FINAL REPORT

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

for

AF- AFOSR [redacted] 414-67

Studies to Improve the Estimation of Seismological Hypocenters and Magnitudes

by

Eugene Herrin
and
William T. Tucker

SMU No. 80-34

Southern Methodist University
Dallas, Texas 75222
Department of Geological Sciences

6 January 1972

Eugene Herrin, Professor
Principal Investigator
A.C. 214-692-2760

BEST
AVAILABLE COPY

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Southern Methodist University
Department of Geological Sciences
Dallas, Texas 75222

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

3. REPORT TITLE

Studies to Improve the Estimation of Seismological Hypocenters and Magnitudes

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Scientific.........Final

5. AUTHOR(S) (First name, middle initial, last name)

Eugene Herrin and William T. Tucker

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
6 January 1972	109	33

8a. CONTRACT OR GRANT NO

AF- AFOSR-414-67

b. PROJECT NO.

8652 and AO 292-78

c.

62701D

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

AFOSR - TR - 72 - 0426

10. DISTRIBUTION STATEMENT

Approved for public release;
distribution unlimited.

11. SUPPLEMENTARY NOTES

TECH, OTHER

12. SPONSORING MILITARY ACTIVITY

Air Force Office of Scientific Research/NPG
1400 Wilson Boulevard
Arlington, VA 22209

13. ABSTRACT

This report covers the work accomplished under AFOSR grant number AFOSR 414-67 which has not yet been covered in our existing reports or publications. In particular, we have attempted to present the theory and empirical results which we believe represent the state-of-the-art in travel-time estimation and the determination of epicenters. Many of the findings follow from results previously reported. These results are not discussed in detail in this report; thus, the reader may need to refer to some of the references in order to follow the discussion. Many of the pertinent reference papers are in the "Special Number-1968 Seismological Tables for P Phases," Bulletin of the Seismological Society of America, Vol. 58, No. 4.

In the report we first develop a useful, mathematical model for the teleseismic travel-times and then discuss the statistical techniques involved in the estimation of model parameters. The problems which limit the accuracy of epicenter determinations are discussed in detail. We believe that the theory and numerical results presented in this report provide a good appriasal of current capabilities for epicenter determination and a satisfactory base for further studies designed to improve the accuracy of estimated location parameters.

DD FORM 1473

FINAL REPORT

Project Code Number: 8652

Name of Grantee: Southern Methodist University

Date of Grant: 1 July 1967

Amount of Grant: $274,356

Grant Number: AF-AFOSR 414-67 Mod. No. 67-0414C

Grant Termination Date: 30 June 1971

Principal Investigator: Eugene Herrin

Title of Work: Studies to Improve the Estimation of Seismological Hypocenters and Magnitudes

CONTENTS

INTRODUCTION

This report covers the work accomplished under AFOSR grant number AFOSR 414-67 which has not yet been covered in our existing reports or publications. In particular, we have attempted to present the theory and empirical results which we believe represent the state-of-the-art in travel-time estimation and the determination of epicenters. Many of the findings follow from results previously reported. These results are not discussed in detail in this report; thus, the reader may need to refer to some of the references in order to follow the discussion. Many of the pertinent reference papers are in the "Special Number-1968 Seismological Tables for P Phases", Bulletin of the Seismological Society of America, Vol. 58, No. 4.

In the report we first develop a useful, mathematical model for the teleseismic travel-times and then discuss the statistical techniques involved in the estimation of model parameters. The problems which limit the accuracy of epicenter determinations are discussed in detail. We believe that the theory and numerical results presented in this report provide a good appraisal of current capabilities for epicenter determination and a satisfactory base for further studies designed to improve the accuracy of estimated location parameters.

SEISMIC TRAVEL-TIME MODELS FOR TELESEISMIC P-WAVES

Basic to any statistical technique is the model used to represent the travel times. This model must adequately explain differences between observed and predicted arrival times. In order to be useful the model must (1) have a reasonable physical basis and (2) be estimable in the statistical sense. The first requirement enables the model to be related to underlying theory and the second enables pertinent parameters to be estimated from observed data. We begin with a complete definition of such a model and then consider approximations needed to make the location problem and the estimation of travel times tractable.

Our general model is:

$$t_{ijk} = O_j + T(\Delta_{ij}, h_j) + (\phi\sigma)_{ij} + \epsilon_{ijk} \quad (1)$$

where i refers to the station,

j refers to the source,

k refers to the observation,

t_{ijk} is the observed arrival time,

O_j is the origin time of the source,

$T(\Delta_{ij}, h_j)$ is the predicted travel time from the <u>jth</u> source to the <u>ith</u> station (from double-entry

tables),

Δ_{ij} is the distance from source to station (in great circle degrees),

h_j is the source depth (in kilometers),

$(\phi 6)_{ij}$ is an interaction time term, and

ϵ_{ijk} is the error term.

In this model ϵ_{ijk} is the error of observation and can be examined by holding i and j fixed and looking at repeated observations (varying k). Work of Freedman (1968, 1966) indicates that this error term can be approximated by a mixture of normal distributions. The principal error components are: (Freedman, 1968)

(1) Miscounts - These occur in multiples of seconds, minutes, quarter hours, or even hours, independent of the errors listed below.

(2) Misidentifications - These arise as a result of poor signal-to-noise ratios. This error and the reading error are dependent variables.

(3) Instrumental errors - These arise from variations in paper speed, clock errors, variations in instrument response, etc.

(4) Reading errors - This is the residual error which would remain even though the above were eliminated.

By employing only the better stations and larger events it is possible to virtually eliminate types (2) and (3). In practice, it can then be assumed that any ϵ_{ijk} consists of components (1) and (4). This implies, unfortunately, that the ϵ_{ijk} are contaminated (Tukey, 1960, 1962). Since contamination, even a very small proportion, may vitiate the usual estimators, data sets must routinely be truncated. With the data truncated for outliers, then the usual least squares techniques are applicable (Tukey, 1960; Dixon, 1953). We will return to this point later. However, with truncation it is possible to assume that:

$$\epsilon_{ijk} \sim \text{(Truncated) NID } (0, \sigma^2) \text{ over all } i, j, k. \qquad (2)$$

Studies employing the above considerations and using repeated explosions at the Nevada test site, indicate that σ^2 is of the order of 0.01 to 0.04 sec^2.

Furthermore, we shall place restrictions on the travel time model based on geophysical considerations. We require, from the concept of reciprocity, that all travel time terms be symmetrical; that is,

$$X_{ij} = X_{ji} \qquad (3)$$

The travel time remains the same if we interchange source and station. In addition, we restrict the value of Δ_{ij} such that,

$$25^\circ \leq \Delta_{ij} \leq 100^\circ.$$

The model given by equation (1) (Model 1) meets the two criteria. Consider a fixed source, j, and station, i. The observed arrival time with repeated events at j consists of an origin time plus a calculated travel time (from double entry tables) plus a correction term for the difference in the true travel time from j to i and the tabled value and finally, plus measurement error. Also with explosion data it is possible to estimate the only important unknown, $(\phi\sigma)_{ij}$. With the assumption of normality the unique minimum variance unbiased estimator of $(\phi\sigma)_{ij}$ is

$$\widehat{(\phi\sigma)}_{ij} = \sum_{k=1}^{n} \left(t_{ijk} - O_{jk} - T(\Delta_{ij}, h_j)\right) / n \tag{4}$$

where i, j are fixed,

n = the number of events at source j, and

O_{jk} is the origin time for the k<u>th</u> event at j.

Since in practice it is difficult to repeat explosions at exactly the same site another way to accomplish the estimation is to have one event and a number of recorders at each station. The minimum variance unbiased estimator is again

given by equation (4) with O_{jk} replaced by O_j. The explosion data from the Nevada test site could reasonably fit either of these situations. In fact the procedure amounts to a calibration of the source-station pair.

In practice, the procedure cannot be carried out since the layout in i and j is very incomplete. Suitable explosions have taken place at only a few locations. Asymmetries in the networks and the scarcity of sources means that $(\phi\sigma)_{ij}$, can be estimated for only a few paths. That is, complete travel time calibration can be obtained for a very restricted set of source-station pairs.

However, Model 1 has a more serious drawback. The travel-time term $T(\Delta_{ij}, h_j)$ can come from any travel-time tables. In fact one could use

$$T(\Delta_{ij}, h_j) \equiv 0$$

for every distance and depth and Model 1 would still apply. Thus, in this respect Model 1 is trivial. The usefulness of the $T(\Delta_{ij}, h_j)$ term is that, hopefully, it will be a function of only the distance from the source to the station and $(\phi\sigma)_{ij}$ will be on the order of a few seconds. Let us now consider only surface events and set

$$T(\Delta_{ij},0) = T(\Delta_{ij}).$$

We wish to define a "world wide average travel-time curve" and consider to what extent this travel time is a function of distance only. Computational techniques exist, Engdahl et al., (1968), by which the surface focus case can be extended to cover nonzero depths.

Suppose that events and stations can be placed at will throughout the world. We restrict stations to be on surface land masses and make the spatial layout in i and j complete with the requirement that all distance ranges appear. Suppose, further, that the measurement error is zero. In this case the model for an observed arrival time becomes:

$$t_{ij(\ell)} = O_j + T(\Delta_{(\ell)}) + (\phi\sigma)_{ij(\ell)} \tag{5}$$

where i refers to the station,

j refers to the source,

(ℓ) is a dummy subscript and refers to the (ℓ)<u>th</u> distance range,

$t_{ij(\ell)}$ is the observed arrival time,

O_j is the origin time of the source,

$T(\Delta_{(\ell)})$ is the predicted travel time from <u>fixed</u> tables,

$\Delta_{(\ell)}$ is the (ℓ)<u>th</u> distance from source j to station i, and

$(\phi\sigma)_{ij(\ell)}$ is an interaction time term.

The interaction term accounts for the failure of the predicted travel time to agree with the observed travel time. The model (Model 2) given by equation (5) describes our population of travel times. That is, Model 2 gives a description of any specific one of the total travel times that we will consider. By indefinitely increasing the number of sources and stations the population approaches the population of all possible travel times from any source to any station, with the restriction that stations are on surface land masses.

Known geological and geophysical properties of the Earth indicate that different regions will have different seismic properties. To the extent possible we, further, assume that the Earth has been stratified according to differing seismic path properties and that an equal number of events occur in each stratum. We stratify primarily to insure that the population gives a truly "world wide"

description and to reduce the sampling variance of any estimators. That stratification can significantly reduce the variance in sampling is well known (Cochran, 1963). We consider only source-station pairs that meet the distance requirement,

$$25^\circ \leq \Delta \leq 100^\circ$$

and suppose that each event is seen by a number of stations, N_j. We do not require that the N_j are all equal. We will consider this further when we discuss randomization. We will assume that all events are in the magnitude range 5½ to 7. We make this assumption in order to keep the observational error of any real data set independent of the randomization errors due to averaging over model effects.

With these restrictions we suppose that a population of source-station pairs and their corresponding travel times is selected and <u>fixed</u>. This population of seismic events then consitutes the population of events for which we will define a world wide average travel time function. Further, any estimators ultimately considered will be

unbiased for this population and, in general, will be biased for any other population. This emphasizes the care that is necessary in actually selecting the population or in its hypothetical selection. For example, in estimating travel time corrections (Tucker, et al., 1968) a world wide population of events was envisioned. This population was suitably stratified and events selected at random from each stratum. We thus believe that the 1968 Herrin travel times (Herrin, et al., 1968) are unbiased for this world wide population. The actual problem is more complex than we here state and we will amplify later. However, our present point is that the sample produces unbiased estimators for the population sampled.

Recalling equation (5), we see that each travel time in the population can be expressed as

$$t_{ij(\ell)} = O_j + T(\Delta_{(\ell)}) + (\phi\sigma)_{ij(\ell)}. \tag{5}$$

Define

$$(\phi\sigma)^*_{ij(\ell)} = (\phi\sigma)_{ij(\ell)} - (\overline{\phi\sigma})_{..(\ell)}, \tag{6}$$

where

$$\overline{(\phi\sigma)}_{..(\ell)} = \sum_{\substack{\text{All } i,j \text{ at} \\ \text{distance } \ell}} (\phi\sigma)_{ij(\ell)} \Big/ N_{..(\ell)} ,$$

and

$N_{..(\ell)}$ = the number of travel paths at distance ℓ.

Also set

$$T^{*}(\Delta_{(\ell)}) = T(\Delta_{(\ell)}) + \overline{(\phi\sigma)}_{..(\ell)} . \tag{7}$$

With these definitions equation (5) can be rewritten as

$$t_{ij(\ell)} = O_j' + T^{*}(\Delta_{(\ell)}) + (\phi\sigma)^{*}_{ij(\ell)} . \tag{8}$$

For future reference observe that

$$(\phi\sigma)^{*}_{..(\ell)} = \sum_{\substack{\text{All } i,j \text{ at} \\ \text{distance } \ell}} (\phi\sigma)^{*}_{ij(\ell)} = 0 \text{ for each } \ell \tag{9}$$

and that $T^{*}(\Delta_{(\ell)})$ is a function of only the distance from the source to the event. In general, $(\phi\sigma)^{*}_{ij(\ell)}$ varies not only with j and i but with ℓ as well. However, in view of (9) this is a local change with ℓ for each j-i pair. The term $T^{*}(\Delta_{(\ell)})$ is the world wide average travel time function and varies only with ℓ (that is with Δ).

It is possible to give a physical interpretation to the terms in equation (8). Suppose that the Earth is a sphere with radius r (r is the radius of a sphere with volume equal to the volume of the Earth). In any actual

computations ellipticity corrections are made so that this assumption is valid. We assume that any seismic path is characterized by a velocity distribution $V(r)$. For a given velocity distribution it is possible to solve appropriate differential equations to determine the path and hence the travel time. In general, the velocity structure of different regions of the Earth differs. However, we make the following assumptions concerning the velocity structure:

(1) The region containing any anomalous velocity structure is confined to the crust and upper mantle. The lateral inhomogeneities which produce these anomalies are confined to the upper 500-700 km of the Earth and below that depth the velocity distribution is spherically symmetric.

(2) The velocity distribution of any region of the Earth is such that for distances in excess of 25° the deepest point of the ray path is below 500-700 km. Thus, the P arrivals that make up the population pass twice through the upper 700 km but most of their ray path is in the spherically symmetrical region of the mantle.

We associate $T^*(\Delta)$ with the spherically symmetrical portion of the path. Here $T^*(\Delta)$ represents a virtual travel time dependent on distance in that the symmetrical portion of the velocity distribution may not extend to the surface. The average over many paths represents an apparent symmetrical distribution that extends to the surface and which would produce $T^*(\Delta)$. Then $(\emptyset\sigma)^*_{ij(\ell)}$ represents deviations from a complete spherically symmetrical distribution. Qualitatively, we see that $(\emptyset\sigma)^*_{ij(\ell)}$ tends to decrease with increasing Δ since the inhomogenous part makes up proportionally less of the total path.

Even if the inhomogeneous layer extended below 700 km, say, included the entire path, the interpretation of $T^*(\Delta)$ as a virtual travel time would still be valid. However, in this case there would be no general similarity in any path in the Earth and the usefullness of $T^*(\Delta)$ would be diminished. In this case the variation in $(\emptyset\sigma)^*_{ij(\ell)}$ would be large and the characterization of a P ray path in terms of epicentral distance would have little meaning.

We will henceforth suppose that the assumptions (1) and (2) are in effect and that $T^*(\Delta)$, for the range of Δ of interest, is of the order of hundreds of seconds and

due primarily to a spherically symmetrical velocity distribution. We assume that a further consequence is that $(\phi\sigma)^*_{ij(\ell)}$ is at most an order of magnitude smaller than $T^*(\Delta)$. In any event, $T^*(\Delta)$ is unique once the population under investigation is fixed. Further, for an increasing coverage of event-station pairs and assumptions (1) and (2) holding the dependancy of $T^*(\Delta)$ on the parent population will decrease.

Equation (8) will be termed Model 3 and represents the population under study. Model 3 with the addition of observation error will be called model 4. Model 4 is not trivial in that $T^*(\Delta)$ is well defined and essentially unique. Thus, Model 4 does not have the serious defect of Model 1 but again, in practice, neither $T^*(\Delta)$ and $(\phi\sigma)^*_{ij(\ell)}$ can be estimated. Suitable explosions have taken place at only a few locations and, therefore, the layout in i and j is very incomplete. We also define Model 5 as Model 4 without the depth restriction. Employing the computational techniques developed by Engdahl, et al., (1968) it is possible to determine a velocity distribution from $T^*(\Delta)$. Once a velocity distribution is established it is possible to obtain $T^*(\Delta,h)$ where $T^*(\Delta,h)$ represents a world wide average travel time for a distance Δ

and with an event at depth h. It is important to note that $T^*(\Delta,h)$ was not developed by "averaging" as was $T^*(\Delta)$. Thus, $T^*(\Delta,h)$ can only be interpreted as a world wide average of a suitable population to the extent that a spherically symmetrical velocity distribution exists. This is so since the computational methods of Engdahl, et al., (1968) employ $T^*(\Delta)$ and its derivative to "strip" the crustal layers from the surface-focus curve and then by inversion to obtain a velocity distribution which is employed in obtaining $T^*(\Delta,h)$ for h nonzero.

In order to go further in the practical analysis of travel times, we must make further simplifying assumptions. Since assumption (1) seems to hold in practice and assumption (2) is not unreasonable, if we consider principally lower order effects, we shall partition the interaction term as follows:

$$(\phi\sigma)_{ij(\ell)} = \phi_{ij(\ell)} + \sigma_{ij(\ell)} \qquad (10)$$

where $\phi_{ij(\ell)}$ is primarily associated with the station and $\sigma_{ij(\ell)}$ is primarily associated with the source. It follows that

$$\overline{(\phi\sigma)}_{..(\ell)} = \bar{\phi}_{..(\ell)} + \bar{\sigma}_{..(\ell)} \qquad (11)$$

where

$$\bar{\phi}_{..(\ell)} = \sum_{\substack{\text{All } i,j \text{ at} \\ \text{distance } \ell}} \phi_{ij(\ell)} \Big/ N_{..(\ell)} \,,$$

$$\bar{\sigma}_{..(\ell)} = \sum_{\substack{\text{All } i,j \text{ at} \\ \text{distance } \ell}} \sigma_{ij(\ell)} \Big/ N_{..(\ell)} \,,$$

and $N_{..(\ell)}$ is as before. Then

$$(\phi\sigma)^{*}_{ij(\ell)} = (\phi\sigma)_{ij(\ell)} - (\overline{\phi\sigma})_{..(\ell)}$$

$$= \phi_{ij(\ell)} + \sigma_{ij(\ell)} - \bar{\phi}_{..(\ell)} - \bar{\sigma}_{..(\ell)}$$

$$= \phi_{ij(\ell)} - \bar{\phi}_{..(\ell)} + \sigma_{ij(\ell)} - \bar{\sigma}_{..(\ell)}$$

$$= \phi^{*}_{ij(\ell)} + \sigma^{*}_{ij(\ell)} \qquad (12)$$

where

$$\phi^{*}_{ij(\ell)} = \phi_{ij(\ell)} - \bar{\phi}_{..(\ell)}$$

and

$$\sigma^{*}_{ij(\ell)} = \sigma_{ij(\ell)} - \bar{\sigma}_{..(\ell)} \,.$$

Again note that

$$\sum_{\text{All } i,\, j \text{ at distance } \ell} \phi^*_{ij(\ell)} = 0 \quad \text{for every } \ell \tag{13}$$

and

$$\sum_{\text{All } i,\, j \text{ at distance } \ell} \sigma^*_{ij(\ell)} = 0 \quad \text{for every } \ell. \tag{14}$$

Our models now become Model 3a, Model 4a and 5a, respectively, and are given by

$$t_{ij(\ell)} = O_j + T^*(\Delta_{(\ell)}) + \phi^*_{ij(\ell)} + \sigma^*_{ij(\ell)} \tag{15}$$

$$t_{ij(\ell)k} = O_j + T^*(\Delta_{(\ell)}) + \phi^*_{ij(\ell)} + \sigma^*_{ij(\ell)} + \epsilon_{ij(\ell)k} \tag{16}$$

and

$$t_{ij(\ell)k} = O_j + T^*(\Delta_{(\ell)}, h_j) + \phi^*_{ij(\ell)} + \sigma^*_{ij(\ell)} + \epsilon_{ij(\ell)k} \, . \tag{17}$$

We note that if k is one, that is, only one event is considered, then it is impossible to separate $\phi^*_{ij(\ell)}$, $\sigma^*_{ij(\ell)}$ and $\epsilon_{ij(\ell)k}$. Thus we must know $\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$, have suitable estimates of these terms or have a simpler model hold if we are to observe variances of the order of 0.01 to 0.04 sec^2.

One possible simplification was made by Cleary and Hales (1966 a), who assumed that the partitioned interaction terms,

$\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$ are independent of distance and azimuth. The resulting "zero th order" model is,

$$t_{ij} = O_j + T^*(\Delta_{ij}, h_j) + \phi_i + \sigma_j + \epsilon_{ij} \tag{18}$$

where

ϕ_i is the "station time term", and

σ_j is the "source time term".

If the origin time is unknown then O_j and σ_j are statistically confounded so that we may write

$$t_{ij} = O_j^* + T^*(\Delta_{ij}, h_j) + \phi_i + \epsilon_{ij} \tag{19}$$

where O_j^* is the "biased" origin time.

It is possible to express the zero th order model terms ϕ_i and σ_j as suitable means of $\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$ and the error term ϵ_{ij} as a sum of $\epsilon_{ij(\ell)k}$ and mean deviations in the $\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$ terms of Model 5a. However, it is impossible to give a randomization interpretation if k is one so that we will not pursue this further. We, thus, interpret ϕ_i and σ_j as true model terms that are not functions of distance. Furthermore, the error term ϵ_{ij} is also assumed to be independent of distance. In fact ϵ_{ij} is just measurement error under the models given by equations (18) and (19).

If the zero th order model is adequate, the sum-square of residuals (suitably normed) resulting from locations of earthquakes used to estimate 0_j^*, $T^*(\Delta, h)$ and $\emptyset_i$ should be comparable to the variance of ϵ_{ij}. In the Cleary-Hales (1966) studies this sum was of the order of 1 sec^2 (Hales, personal communication, 1969), much larger than expected from the simplified travel time model (less than 0.1 sec^2). Clearly the error term, ϵ_{ij}, in equation (19) was absorbing more "error" than would be expected due to uncertainties in measurement and timing alone. The Cleary-Hales studies indicate that the simplified model is inadequate in explaining the interaction terms of Models 3a, 4a and 5a.

Later Douglas and Lilwall (1968) made a further model simplification in studying essentially four events in the Pacific whose locations were known from independent studies. Douglas and Lilwall in effect assume that σ_j is constant for the four event areas. The data are from the Eniwetok series, Bikini series, a large Hawaiian earthquake and Longshot. The standard errors of observation as quoted by Douglas and Lilwall are Longshot: 0.8 sec, Hawaii: 1.2 sec, Eniwetok: 0.5 sec (average) and Bikini: 0.5 sec (average). The range on Eniwetok is 0.3-0.7 sec

and on Bikini 0.4-0.6 sec. Thus it again appears that the error term in a simplified model is absorbing more "error" than would be expected. In this model the station time terms, ϕ_i, are independent of source location and could thus be equally well determined (i.e. unbiased estimates obtained) from any given explosion. Joint estimation using four events serves only to reduce the variance of the estimates under the assumed model.

It follows as an implication of this model that estimates of station time terms obtained in the manner just described should be equally applicable to sources anywhere. Ignoring a base line shift (which could be due to not estimating the constant source term) the 1968 Douglas-Lilwall estimates differ from those given by Lilwall and Douglas (1970). A further consequence of this model is that the estimated time terms should never appear to be a function of source location. That is, time terms must not appear to vary with distance or azimuth from station to epicenter. Cleary and Hales (1966b) later reported that in fact "the remaining residuals were azimuthally dependent for some stations". Also Lilwall and Douglas (1970) obtained azimuthally dependent estimates which for many stations are statistically significant. Thus, the simplified model

and on Bikini 0.4-0.6 sec. Thus it again appears that the error term in a simplified model is absorbing more "error" than would be expected. In this model the station time terms, ϕ_i, are independent of source location and could thus be equally well determined (i.e. unbiased estimates obtained) from any given explosion. Joint estimation using four events serves only to reduce the variance of the estimates under the assumed model.

It follows as an implication of this model that estimates of station time terms obtained in the manner just described should be equally applicable to sources anywhere. Ignoring a base line shift (which could be due to not estimating the constant source term) the 1968 Douglas-Lilwall estimates differ from those given by Lilwall and Douglas (1970). A further consequence of this model is that the estimated time terms should never appear to be a function of source location. That is, time terms must not appear to vary with distance or azimuth from station to epicenter. Cleary and Hales (1966b) later reported that in fact "the remaining residuals were azimuthally dependent for some stations". Also Lilwall and Douglas (1970) obtained azimuthally dependent estimates which for many stations are statistically significant. Thus, the simplified model

appears to be inadequate in that the interaction terms, $\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$, can not adequately be represented by two simple time terms.

Cleary and Hales (1966b) and Herrin and Taggart (1968a) have, using statistical analyses of earthquake data, shown that the station time terms for a number of stations are dependent upon azimuth from station to source. These findings, however, may not be, strictly speaking, pertinent to this discussion in that they are based on a simplified travel time model. To use these results as evidence for keeping the interaction term in the travel time model is, in a sense, begging the question. However, there are other studies of azimuthal effects which do not suffer from this drawback. Ritsema (1959) showed that normalized amplitudes of P-waves recorded at Djakarta varied sinusoidally with azimuth. Using P and pP phases recorded by the Berkely network, Otsuka (1966) found the time terms to be sinusoidal functions of azimuth to the event. Nizai (1966) reported the same effect at the Tonto Forest Array. Bolt and Nuttli (1966) studying data from the Berkely network showed conclusively that the time terms for Shasta and Marysville varied sinusoidally with

azimuth. The amplitude of the variation was greater than 1 sec. Errors in epicentral determinations and in the travel time tables used enter only as second order effects in these studies.

Reciprocity requires that source as well as station time terms must be considered, in principle, to vary with azimuth. If an explosion were fired beneath the Shasta station we would expect to see a variation of travel time with azimuth of at least ± 1 sec at teleseismic distances, Thus we are forced to return to the more complicated travel time model,

$$t_{ij(\ell)k} = O_j + T^*(\Delta_{(\ell)}, h_j) + \phi^*_{ij(\ell)} + \sigma^*_{ij(\ell)} + \epsilon_{ij(\ell)k} \quad (17)$$

With mild regularity assumptions we can expand the two interaction terms as Fourier series in azimuth angle with the coefficients being functions of epicentral distance. Thus,

$$\phi^*_{ij(\ell)} = \phi_{i\ell} + {}_1A_{i\ell} \sin Z_{ij} + {}_1B_{i\ell} \cos Z_{ij} + {}_2A_{i\ell} \sin 2Z_{ij} + {}_2B_{i\ell} \cos Z_{ij} + \ldots$$

and

$$\sigma^*_{ij(\ell)} = \sigma_{j\ell} + {}_1A_{j\ell} \sin Z_{ji} + {}_1B_{j\ell} \cos Z_{ji} + {}_2A_{j\ell} \sin 2Z_{ji} + {}_2B_{j\ell} \cos Z_{ji} + \ldots$$

where "ℓ" indicates the dependence of a parameter on epicentral distance, $\Delta_{(\ell)}$.

We now write a first order model, neglecting second and higher order terms in the Fourier expansion, as follows:

$$t_{ij(\ell)} = O_j + T^*(\Delta_{(\ell)}, h_j) + \phi_{i\ell} + A_{i\ell}\sin Z_{ij} + B_{i\ell}\cos Z_{ij} + \sigma_{j\ell} + A_{j\ell}\sin Z_{ji} + B_{j\ell}\cos Z_{ji} + \epsilon^*_{ij(\ell)} \quad (20)$$

where Z_{ij} is the azimuth angle from station to source, $Z_{j\ell}$ is the angle from source to station and $\epsilon^*_{ij(\ell)}$ is $\epsilon_{ij(\ell)}$ plus the remaining Fourier terms of the expansion of $\phi^*_{ij(\ell)}$ and $\sigma^*_{ij(\ell)}$. Note that the symmetry of doubly indexed variables has been lost in this expression, not because of a failure of reciprocity, but because the two azimuth angles associated with the ends of a great circle arc are not simply related. This travel time model is an expansion to the first order in azimuth angle, of the complete interaction model (Model 5) under the assumption that the interaction term can be separated into a source effect and a station effect.

Using the explosion data where O_j, $\Delta_{(\ell)}$ and h_j are known, we can attempt to estimate the parameters in the first

order model, provided sources and receivers are well distributed geographically and in distance and azimuth relative to one another. Attempts to make such estimates have generally failed because known sites of large explosions are rare.

A further simplification can be obtained if the distance dependence of the Fourier coefficients is neglected. As final model forms we consider, Model 6,

$$\begin{aligned} t_{ij(\ell)} = O_j + T^*(\Delta_{(\ell)}) + \phi_i + A_i \sin Z_{ij} \\ + B_i \cos Z_{ij} + \sigma_j + a_j \sin Z_{ji} \\ + b_j \cos Z_{ji} + \epsilon^{**}_{ij(\ell)} \end{aligned} \tag{21}$$

and, Model 7,

$$\begin{aligned} t_{ij(\ell)} = O_j + T^*(\Delta_{(\ell)}, h_j) + \phi_i + A_i \sin Z_{ij} \\ + B_i \cos Z_{ij} + \sigma_j + a_j \sin Z_{ji} \\ + b_j \cos Z_{ji} + \epsilon^{**}_{ij(\ell)} \end{aligned} \tag{22}$$

where $\epsilon^{**}_{ij(\ell)}$ is $\epsilon_{ij(\ell)}$ plus the difference between $\phi^*_{ij(\ell)}$ and $\phi_i + A_i \sin Z_{ij} + B_i \cos Z_{ij}$ and the difference between $\sigma^*_{ij(l)}$ and $\sigma_j + a_j \sin Z_{ji} + b_j \cos Z_{ji}$.

Models 6 and 7 represent expansions to the first order in azimuth angle with constant coefficients under the assumption that the interaction term can be separated into a source effect and a station effect.

STATISTICAL PROBLEMS IN THE ESTIMATION OF MODEL PARAMETERS

With the exception of measurement errors, the population described by Model 5a is deterministic; that is, repeated observations on a fixed travel-time path will result in essentially the same travel time. Any simplified model, for example, Model 7, will thus have a fixed error term that represents the difference in the true travel time and that predicted from the simplified model. In an attempt to treat this fixed error term we assume that any seismic event can be considered to have been selected at random from the population of possible events. Observe that if we do not employ randomization that this model difference remains a fixed and unknown time error. Hence any estimation techniques must be arbitrary. To avoid this dilemma we chose to employ randomization.

Consistent with our model studies we suppose that a fixed set of events and stations has been determined. How representative of the whole Earth our final results are depends on the population of source-station pairs. Thus the 1968 P travel times may be biased toward the seismically active regions of the Earth since approximately 98% of the data

employed came from these regions. Also the results may be biased toward earthquakes since only 13 of the total of 292 events employed were shots. That is, we cannot reasonably assert that the sample employed in the 1968 P travel time study (Special Number - 1968 Seismological Tables for P Phases, Bull. Seism. Soc. Am. 58 No. 4) was a random sample from a world wide population or a population with extensive shot data. However, if we restrict the scope of the population then we can claim that we obtained a random sample. It is to this restricted population that we make inferences.

Until data are obtained in the less active regions of the Earth there is no way to assertain the usefulness of the 1968 P travel times for use in these areas. However, in view of the known geological and geophysical properties of the Earth, we believe that the 1968 times should be quite useful; recall that a spherically symmetric velocity distribution appears to hold for the Earth. Thus we believe that the restricted population is representative of the whole Earth as regards earthquakes.

We could similarly argue that the results would apply to shots in general. In view of independent work by Veith and Clawson (1972) we shall consider the shot data problem somewhat

differently. Figure 1 and Table 1 are excerpted from Veith and Clawson (1972) which the authors were kind enough to let us use before publication. Figure 1 contains a plot of the residuals from the 1968 P times of the shot data given in Table 1. In preparing Figure 1 multiple observations of travel times along the same path were averaged and considered as one. Also any mean source effect was removed from the data on each event before plotting in Figure 1. The removal of the mean source effect will reduce the variability in the data but should not vitiate the results if most sources see a wide range of distances. This is the case (Clawson, personal communication, 1971). Furthermore, no station corrections were applied. Thus Figure 1 presents the residuals from the 1968 P times in a set of 43 shots with 2044 observations at teleseismic distances. The data set should not be biased due to path overloading or mean removal. The results are most pleasing. Visual inspection reveals that there is no appreciable mean shift from the zero axis beyond 25°.

In order to quantitatively investigate the data the actual residuals were read from Figure 1 by overlaying graph paper. The readings were made to .1 second. The results of anyalyzing this data on a 5° cell basis are presented in

Table 2. Due to the close agreement of the computed standard deviations with the plotted curve calculated on a 2° basis we believe not much accuracy was lost. Again the results are pleasing. In the distance range 25° - 100° the largest mean is .392 sec. at 35° - 40°. The largest relative error is approximately one part per thousand and occurs in the 30° - 35° and 35° - 40° cells. These errors are so slight that they would never affect a seismic location with reasonable station coverage. In fact if one is only concerned with location then the removal of the source mean from the residual data has no effect since for adequate networks this could only bias the origin time estimate (Tucker et al., 1968).

If we assume that the observations are independent normally distributed, the efficacy of which can be questioned, then we can make t tests for each cell. The t values are in the last column of Table 2. If we choose an α of .01, then the means of cells 25°-30°, 30°-35° and 95°-100° are significantly different from zero. The degrees of freedom for any of these tests are so large that they are effectively infinite. The critical t value is 2.576 for a two-sided test. Thus only the 35°-40° cell appears to be definitely significant. However, since any

travel-time curve must possess certain differential properties we believe a better procedure is to fit a smooth curve to the data. With randomization we can obtain "optimal" estimates of the fitted curve without the assumptions of independence and normality. We return to these considerations after we finish the discussion of the randomization procedure.

We thus believe that the 1968 P times are not seriously biased against the shot data population of the Veith-Clawson work. Again geological and geophysical considerations lead us to believe that the restricted population is representative of the whole Earth. We conclude that while we sampled a restricted population in the 1968 studies that this population is representative of the whole Earth with respect to both earthquakes and shots. We will consider this point again in the section on estimation.

However, representative or not, we envision a final population of event-station pairs and their corresponding travel times. We list these events as shown in Figure 2 and suppose that there are K strata, M events and N stations in the population. Not all stations see all events, that is, are within the distance restrictions. Also each event-station pair is at some unique distance value. We suppose that the

range of Δ has been broken into cells and that there are L of these. We assume that the cells are of equal width and small enough that the effects of grouping are negligible. The events are classified first according to stratum and second according to event within a stratum, that is, all events are different so that event 1 in stratum 1 is different from event 1 in any other stratum and so on. However, the N stations under each individual event are the same N stations, that is, station 1 in stratum 1 - event 1 is the same station 1 in all other stratum-event combinations. Since each event-station pair has a unique distance there is one travel time value for each row of the figure. These are indicated by the x's in the body of the figure. The randomization scheme is given by the following procedure:

(1) Select by simple random sampling, that is, with equal probabilities and without replacement, k_1 of the K strata where $1 \leq k_1 \leq K$.

(2) Select by simple random sampling m_1 of the M events in each stratum selected in (1) where $1 \leq m_1 \leq M$.

(3) Let N_{km} be the number of stations that see the $\underline{m\text{th}}$ event in the $\underline{k\text{th}}$ stratum. Select by simple random sampling n_{km} of the N_{km} stations for each event selected in (2) where

$1 \leq n_{km} \leq N_{km}$.

In the 1968 studies steps (1) and (3) were carried out (Tucker, et al., 1968 and Herrin, et al., 1968). The available data were grouped on a 1° latitude by 1° longitude basis and the largest source event selected from each group. The groups are the strata and $k_1 = K$. We assume that the largest event in a stratum does not behave systematically in the population so that the selction of events in the strata can be thought of as random. Here $m = 1$ and each k-m combination is equivalent to a particular j in our previous notation. Step (3) was met in that $n_{km} = N_{km}$. The most serious difficulty with the 1968 procedure is that the cell grouping will not produce strata with equal sizes. However, since most of the data are from active seismic regions the assumption of equal size strata should be adequate. Note that since $n_{km} = N_{km}$ we have $n_{km}/N_{km}=1$.

Now consider the general randomization scheme with the restriction that $n_{km}/N_{km}= \alpha$, a constant, not necessarily equal to one. Then one can show that the probability that an individual event-station pair is included in a sample is $(k_1/K)(m_1/M)\alpha$ (Cochran, 1963). Thus the inclusion probabilities are equal. If we consider a given distance range, then the number of observations in

the whole sample at this range is a random variable. Similarly the number of observations made by any station is random. Thus it is not obvious that the usual statistics (sample mean, sample variance, regression estimates, etc.) are applicable.

We shall now discuss a general estimation scheme that will justify the use of the ordinary statistics and indicate some further implications of randomization. The technique is a generalization of the linear regression estimator in sampling theory (Cochran, 1963). Suppose we have a population of observables Y_i, $i = 1, 2, \ldots, N$. We would like to predict the Y_i values in terms of a set of independent variables $X_{1i}, X_{2i}, \ldots, X_{ki}$, that is, for given i and hence $X_{1i}, \ldots, X_{ki}$ we desire an estimate of Y_i. In order to obtain a solution we assume that the Y's can be suitably estimated by linear combinations of known functions of the X's, that is, we assume a set of basis functions exists. We employ these basis functions in only finite sums so that we, in general, cannot approximate the Y's exactly. We write

$$Y_i = \beta_0 + \beta_1 h_1(X_{1i}, \ldots, X_{ki}) + \cdots + \beta_n h_n(X_{1i}, \ldots, X_{ki}) + \delta_i \quad (23)$$

where

$$\delta_i = Y_i - [\beta_0 + \beta_1 h_1(X_{1i}, \cdots, X_{ki}) + \cdots + \beta_n h_n(X_{1i}, \cdots, X_{ki})].$$

we suppose that the β's are such that

$$Q(\beta) = \sum_{i=1}^{N} \delta_i^2$$

is a minimum. Thus the (population) predictor is determined by least squares.

This definition of the predictor is admittedly arbitrary but can be justified as follows. Suppose that an individual Y_i is selected at random from the population. Then one by randomization theory (Kempthorne, 1955; Cornfield, 1944) can show that if we express this Y_i as in equation (23) with

$$\mu_i = \beta_0 + \beta_1 h_1(X_{1i}, \cdots, X_{ki}) + \cdots + \beta_n h_n(X_{1i}, \cdots, X_{ki}) \quad (24)$$

then Y_i becomes a random variable with mean μ_i and can be expressed as

$$Y_i = \mu_i + e_i. \quad (25)$$

The mean of e_i is zero and the variance of e_i is

$$V(e_i) = \sum_{i=1}^{N} \delta_i^2 / N. \quad (26)$$

We see from equation (26) that $V(e_i)$ is a minimum consistent with our model requirements, that is, linear basis. If no Y_i is any more important than any other than our assumptions on the β's is not unreasonable.

Since the population predictor is based on linear least squares we have an immediate solution for the β's (Graybill, 1961);

$$\beta = (X'X)^{-1}X'Y, \tag{27}$$

where

$$X = \begin{bmatrix} 1 & h_1(X_{11},\cdots,X_{k1}) & \cdots & h_n(X_{11},\cdots,X_{k1}) \\ 1 & h_1(X_{12},\cdots,X_{k2}) & \cdots & h_n(X_{12},\cdots,X_{k2}) \\ \vdots & \vdots & & \vdots \\ 1 & h_1(X_{1N},\cdots,X_{kN}) & \cdots & h_n(X_{1N},\cdots,X_{kN}) \end{bmatrix},$$

$$\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}, \quad Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{pmatrix}$$

and $^{-1}$ denotes inverse and $'$ denotes transpose. Thus the β's are linear functions of the population Y's. In practice it is often impossible to determine the β's since we do not know the Y's. Thus we must estimate the β's by some scheme.

Since the β's are linear in the Y's, if we take a simple random sample, then an optimum estimate of X'Y is given by

$$\widehat{X'Y} = \frac{N}{m} X_m' Y_m$$

where

$$X_m = \left[\text{Sampled rows of } X\right],$$

$$Y_m = \left(\text{Sampled rows of } Y\right),$$

and m is the sample size, that is, X_m is the matrix of sampled basis functions and Y_m is the corresponding sampled Y values.

The estimator $\widehat{X'Y}$ is called a Horvitz-Thompson (HT) estimator and is only an optimum estimator since no uniformly unbiased minimum variance estimator of the totals X'Y exists when sampling from finite populations (Godambe, 1955; Hanurav, 1968). Relaxing the uniformity criterion Hanurav (1968) has shown that the HT estimator does have a certain optimal property and is unique. The HT estimator is unbiased. For its complete properties see Hanurav (1968).

Thus

$$E(\widehat{X'Y}) = X'Y$$

so that

$$E\left[(X'X)^{-1}(\widehat{X'Y})\right] = (X'X)^{-1}X'Y$$
$$= \beta$$

as we see from equation (27). Thus an optimal estimator of β is given by

$$\hat{\beta} = (X'X)^{-1}\left(\frac{N}{m}\right) X_m' Y_m \ . \qquad (28)$$

In order to employ equation (28) we must know N and X. In the 1968 studies these quantities were not known. However, we have

$$(X'X)^{-1}\left(\frac{N}{m}\right) = \left(\frac{m}{N} X'X\right)^{-1}$$

$$= \begin{bmatrix} m & \frac{m}{N}\sum_i h_1(i) & \cdots & \frac{m}{N}\sum_i h_n(i) \\ \frac{m}{N}\sum_i h_1(i) & \frac{m}{N}\sum_i h_1^2(i) & \cdots & \frac{m}{N}\sum_i h_1(i)h_n(i) \\ \vdots & \vdots & & \vdots \\ \frac{m}{N}\sum_i h_n(i) & \frac{m}{N}\sum_i h_n(i)h_1(i) & \cdots & \frac{m}{N}\sum_i h_n^2(i) \end{bmatrix}$$

where $h_j(i) = h_j(x_{1i}, \ldots, x_{ki})$. If the sample chosen is representative of the population X's, then

$$X_m' X_m \doteq \frac{m}{N} X' X \tag{29}$$

In fact, if the strata are such that the X's are grouped into m equal size sets of common values and we sample such that one from each common set is chosen, then

$$X_m' X_m = \frac{m}{N} X' X$$

as one can easily show. Unfortunately this was only approximately done in the 1968 studies since not all events see all distance ranges and there was no stratification with respect to stations.

However replacing $(^m/N\ X'X)^{-1}$ by $(X'_m\ X_m)^{-1}$ is analogous to forming a linear regression estimator (Cochran, 1963). We thus consider

$$\hat{\beta} = (X_m' X_m)^{-1} X_m' Y \tag{30}$$

which is the usual linear least squares estimator (Graybill, 1961).

We will now develop a bias formula for the general estimator given by equation (30). From equation (25) we have

$$Y_m = X_m \beta + e_m \tag{31}$$

where

$$e_m = (\text{Sampled } e\text{'s}).$$

Thus

$$\begin{aligned}\hat{\beta} &= (X_m' X_m)^{-1} X_m' (X_m \beta + e_m) \\ &= \beta + (X_m' X_m)^{-1} X_m' e_m. \end{aligned} \qquad (32)$$

Now

$$E(X_m' e_m) = \frac{m}{N} X' e \qquad (33)$$

where

$$e = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{pmatrix}.$$

But from equation (25)

$$\begin{aligned} X'e &= X'(Y - X\beta) \\ &= X'Y - (X'X)\beta \\ &= X'Y - X'Y \\ &= \Phi \end{aligned} \qquad (34)$$

Thus, in repeated sampling with fixed size m, $X'_m e_m$ is a random vector distributed about a zero mean vector and for large m the bias should be small.

We have from equation (32) that

$$\hat{\beta} - \beta = (X_m' X_m)^{-1} X_m' e_m$$

$$= \left(\frac{1}{m} X_m' X_m\right)^{-1} \left(\frac{1}{m}\right) X_m' e_m . \tag{35}$$

Now

$$\left(\frac{1}{m} X_m' X_m\right)^{-1} = \left(\frac{1}{N} X'X\right)^{-1} \left(I - \left(\frac{1}{N} X'X - \frac{1}{m} X_m' X_m\right)\left(\frac{1}{N} X'X\right)^{-1}\right)^{-1}$$

$$\doteq \left(\frac{1}{N} X'X\right)^{-1} \left(I + \left(\frac{1}{N} X'X - \frac{1}{m} X_m' X_m\right)\left(\frac{1}{N} X'X\right)^{-1}\right) \tag{36}$$

if $\frac{1}{m} X_m' X_m \to \frac{1}{N} X'X$ in probability and m is suitably large. From equation (36) we obtain, neglecting higher order terms,

$$E(\hat{\beta} - \beta) = \frac{1}{m^2} \left(\frac{1}{N} X'X\right)^{-1} E\, X_m' X_m \left(\frac{1}{N} X'X\right)^{-1} X_m' e_m . \tag{37}$$

Let the entries in any X matrix be denoted by x_{ij} and the entries in $(X'X/N)^{-1}$ be denoted by $\tilde{x}_{ij}$. Set p=n + 1. Then

$$X_m' X_m \left(\frac{1}{N} X'X\right)^{-1} X_m' e_m$$

$$= \begin{pmatrix} \sum_{l=1}^{p} \sum_{k=1}^{m} \sum_{j=1}^{p} \sum_{i=1}^{m} x_{k1} x_{kl} \tilde{x}_{lj} x_{ij} e_i \\ \vdots \\ \sum_{l=1}^{p} \sum_{k=1}^{m} \sum_{j=1}^{p} \sum_{i=1}^{m} x_{kp} x_{kl} \tilde{x}_{lj} x_{ij} e_i \end{pmatrix}$$

as one may show by straight forward but tedious algebra.

We have for the qth entry in equation (38) that

$$E\left(\sum_{\ell=1}^{p}\sum_{k=1}^{m}\sum_{j=1}^{p}\sum_{i=1}^{m} x_{kq}x_{k\ell}\tilde{X}_{\ell j}x_{ij}e_i\right)$$

$$=E\left(\sum_{\ell=1}^{p}\sum_{j=1}^{p}\tilde{X}_{\ell j}\sum_{k=1}^{m}\sum_{i=1}^{m} x_{kq}x_{k\ell}x_{ij}e_i\right)$$

$$=\sum_{\ell=1}^{p}\sum_{j=1}^{p}\tilde{X}_{\ell j}\,E\left(\sum_{i=1}^{m} x_{iq}x_{i\ell}x_{ij}e_i+\sum_{\substack{k=1\\k\neq i}}^{m}\sum_{i=1}^{m} x_{kq}x_{k\ell}x_{ij}e_i\right)$$

$$=\sum_{\ell=1}^{p}\sum_{j=1}^{p}\tilde{X}_{\ell j}\left(\frac{m}{N}\sum_{i=1}^{N} x_{iq}x_{i\ell}x_{ij}e_i+\frac{m(m-1)}{N(N-1)}\sum_{\substack{k=1\\k\neq i}}^{N}\sum_{i=1}^{N} x_{kq}x_{k\ell}x_{ij}e_i\right)$$

$$=\sum_{\ell=1}^{p}\sum_{j=1}^{p}\tilde{X}_{\ell j}\left[\frac{m}{N}\left(\frac{N-m}{N-1}\right)\sum_{i=1}^{N} x_{iq}x_{i\ell}x_{ij}e_i+\frac{m(m-1)}{N(N-1)}\sum_{k=1}^{N} x_{kq}x_{k\ell}\sum_{i=1}^{N} x_{ij}e_i\right]$$

$$=\sum_{\ell=1}^{p}\sum_{j=1}^{p}\tilde{X}_{\ell j}\,\frac{m}{N}\left(\frac{N-m}{N-1}\right)\sum_{i=1}^{N} x_{iq}x_{i\ell}x_{ij}e_i$$

from equation (34). Thus

$$E\left(\sum_{l=1}^{p}\sum_{k=1}^{m}\sum_{j=1}^{p}\sum_{i=1}^{m} x_{kq}x_{kl}\tilde{X}_{lj}x_{ij}e_i\right)=m(1-f)S_q \quad (39)$$

where $\qquad f = m/N$

and
$$S_q=\sum_{l=1}^{p}\sum_{j=1}^{p}\sum_{i=1}^{N}\tilde{X}_{lj}x_{iq}x_{il}x_{ij}e_i\Big/(N-1)\,. \quad (40)$$

Finally, from equation (37) we have that

$$E(\beta-\hat{\beta})=\frac{(1-f)}{m}\left(\sum_{k=1}^{p}\tilde{X}_{qk}S_k\right). \quad (41)$$

Therefore the bias is of order 1/m if S_k is of order unity for every k. A sufficient condition for this is that the x_{ij}'s and Y_i's be bounded. This is the case for the data of the 1968 studies. Thus the bias due to randomization is of order 1/m for those studies.

Following Cochran (1963 pages 194-196) one can show that an assymptotically valid estimator of $\sum_{i=1}^{N}\delta_i^2$ is

$$\hat{V}(e_i)=Y_m'\left[I-X_m(X_m'X_m)^{-1}X_m'\right]Y_m\Big/(m-k) \quad (42)$$

the usual estimator. Recalling equation (30) we see that the usual least squares estimators are appropriate with our randomization model.

Thus employing the usual least squares estimators should produce negligible bias in the travel time estimates since in the distance range 25° - 100° the fewest observations were 81 in the 100° cell. The case for station terms is not as conclusive since we reported estimates for stations with as few as 10 observations. However, most of the station terms were estimated with much more data.

While we have assumed simple random sampling in proving these results, the HT estimator is optimum under general randomization. Also, if the design is balanced so that the inclusion probabilities are equal and further the second order inclusion probabilities are equal, then the bias is again of order $1/m$. We have previously shown that the (first order) inclusion probabilities are equal. Since we sample all strata and use all stations that see any event the second order inclusion probabilities are either 1, $m(m-1)/M(M-1)$ or $(m/M)^2$. We obtain 1 when we consider two stations that see the same event, $m(m-1)/M(M-1)$ when

we consider two events in the same stratum and $(m/M)^2$ when we consider two events in different strata. Thus the second order inclusion probabilities are not equal. However, from the derivation of equation (39) it is obvious that with these inclusion probabilities again the bias is of order 1/m. Also with a sufficiently large total sample size the usual variance estimator is also applicable. Thus these results apply in the randomization scheme employed in the 1968 studies and the usual estimators are applicable in that work.

Let us now consider the general case in which measurement error is also present. Thus suppose that

$$Y_m = X_m \beta + e_m + \epsilon \tag{43}$$

where

$$\epsilon \sim N(\phi, \sigma^2 I).$$

Then

$$E(\hat{\beta}) = E(X_m' X_m)^{-1} X_m' [X_m \beta + e_m + \epsilon]$$
$$= \beta + (X_m' X_m)^{-1} X_m e_m$$

which is the result given earlier in equation (32).

Also

$$E\left(Y_m'\left[I-X_m(X_m'X_m)^{-1}X_m'\right]Y_m\big/(m-k)\right)$$

$$=\frac{1}{(m-k)}\left\{E(X_m\beta+e_m)'\left[I-X_m(X_m'X_m)^{-1}X_m'\right](X_m\beta+e_m)\right.$$

$$\left.+\,E\,\epsilon'\left[I-X_m(X_m'X_m)^{-1}X_m'\right]\epsilon\right\}$$

since ϵ and e_m are independent. Now from equation (42) we have that

$$E\,(X_m\beta+e_m)'\left[I-X_m(X_m'X_m)^{-1}X_m'\right](X_m\beta+e_m)\big/(m-k)$$

$$\doteq V(e_i).$$

Also it is well known (Graybill, 1963) that

$$E\,\epsilon'\left[I-X_m(X_m'X_m)^{-1}X_m'\right]\epsilon\big/(m-k)$$

$$=\sigma^2.$$

Thus

$$E\left(Y_m'\left[I-X_m(X_m'X_m)^{-1}X_m'\right]Y_m\big/(m-k)\right)$$

$$\doteq V(e_i)+\sigma^2$$

and again the usual estimator is appropriate. Finally, in the general case the usual least squares estimators are assymptotically unbiased. If we require that the regression estimator be optimal with respect to the randomization scheme then the derived estimator is the only one since the HT estimator is unique.

We now consider one final result of randomization. Suppose that a single event is drawn at random; that is, one stratum is selected at random and then one event is selected at random within the selected stratum. We suppose that Model 4 holds and $T^*(\Delta)$ is known. With randomization $(\phi\sigma)^*_{ij(\ell)}$ becomes random. Also suppose that for each event in the population each station is at a different distance cell. Then for a fixed distance, ℓ, we have

$$E\left(t_{ij(\ell)} - T^*(\Delta_{(\ell)})\right) = E\left((\phi\sigma)^*_{ij(\ell)} + \epsilon\right)$$
$$= \sum_{\text{All } i,j \text{ at } \ell} (\phi\sigma)^*_{ij(\ell)}$$
$$= 0$$

from equation (9). Thus, if we draw random samples from our population, we expect that at each distance, ℓ, the residuals from $T^*(\Delta)$ would average out to zero. This is essentially the case with the Veith-Clawson data. It is in this sense that a world-wide average travel-time function is unambiguously defined.

Note further that if an event is fixed then the $(\phi\sigma)^*_{ij(\ell)}$ are not random. The only random quantity is ϵ. If one is then attempting to locate this fixed event with

a fixed set of stations, then the location may be significantly biased. Longshot is an example of this phenomenon. If repeated shots were made on Amchitka and a given set of stations used to locate these shots, then the 1968 travel times would, in general, give locations approximately 25 km to the north. One could argue that the 1968 times are not "tuned" to Amchitka and this is the case. However, if the 1968 times are altered to fit Amchitka, the altered times would not fit the Veith-Clawson data. We shall return to this point when we consider possible models of the Veith-Clawson data. This being the case, it should be obvious that how well a given world-wide average travel-time function locates a specific event is not a suitable criterion of the validity of the travel-time function.

The problem now arises of how to make statistical tests with the randomization scheme employed. Three approaches are possible. One is to consider non-parametric techniques. We rule these out since ranks, runs, etc. of travel times make little sense. A second possibility is to employ asymptotic theory to justify the normality assumption. In general, this requires conditions on the higher order moments of the population (Madow, 1948).

In our problem it is impossible to verify these conditions. However, with "reasonable" sample sizes for many problems the assumption of normality is adequate (Kempthorne, 1955). Thus we believe the assumption of normality is justified in the 1968 studies for the estimation of travel times. This should also be true for the estimation of station corrections for those stations that had large numbers of observations. A third approach is to employ tests that are robust; that is, perform adequately when the assumption of normality does not hold. There is a vast literature on the subject of robustness (Govindarajulu and Leslie, 1970). Much of this work is on specific problems and employs Monte Carlo studies. The results vary. In some cases the usual test procedures, t, F, analysis of variance, etc., work well and in some they do not. Now not only sample size is important, but also population size since randomization produces correlation and the standard assumption is of independent normal observations. Since there exist broad classes of data for which the usual tests are robust (with suitable sample and population size) we again believe that the usual test procedures are adequate in the 1968 studies in those cases where the sample size was large.

To compliment these considerations histograms of the travel-time residuals were plotted for each distance cell and for ten selected stations. None of the plots exhibited extreme non-normality. We made no goodness-of-fit tests since these must necessarily be chi-square tests (population parameters were not known).

With respect to the Veith-Clawson data, we can again expect the usual tests to be robust since the fewest number of observations, 50, occurred in the 95°-100° cell. Thus the t-test indicates that at a significance level of .01 that the mean of cell 35°-40° is not zero. However, as stated earlier, any travel-time function must possess certain differential properties, and the t-test does not indicate how this can be done. Further, for these data, the t-test is not exactly correct since the data are corrected residuals and hence correlated at each distance, l. Since optimum estimators can be obtained without the assumption of normality, we shall do this and, in any event, we will have valid estimates.

We suppose that the Veith-Clawson data constitute a random sample from a suitable population. Due to the path averaging this assumption should hold approximately. The

fact that no Russian or Chinese stations are represented in the data, and all events are in the northern hemisphere implies that the population is even more restricted than that of the 1968 studies. These data represent the most extensive collection of teleseismic shot data analyzed. In particular it includes all shots employed by Lilwall and Douglas (1970), Cleary and Muirhead (1969), and Muirhead and Cleary (1969). These works are especially important since the results indicate that the 1968 times are in error. Ignoring a baseline shift the corrections with respect to these works are (Douglas, 1970):

$$T_{LD} = T_{1968} - 0.006\Delta$$

and

$$T_{CM} = T_{1968} - 0.023\Delta$$

where LD indicates Lilwall-Douglas and CM indicates Cleary-Muirhead and T denotes travel time.

We shall follow these authors and suppose that any error in the 1968 times is given by a linear function, say, $a + b\Delta_{\ell}$. In order to make comparisons with the Cleary-Muirhead results, we shall consider the distance range 20°-100°. Also, we must allow for a mean source correction. Our model is then

$$t_{ij(\ell)} = O_j + T(\Delta_{(\ell)}) + a + b\Delta_\ell + \phi_j + ((\phi\sigma)^*_{ij(\ell)} - \phi_j) + \epsilon_{ij(\ell)} \tag{44}$$

where

$$T^*(\Delta_{(\ell)}) = T(\Delta_{(\ell)}) + a + b\Delta_\ell$$

$$T(\Delta_{(\ell)}) = \text{the 1968 P travel times}$$

and

$$\phi_j = \text{a source time correction term.}$$

Let

$$\epsilon^*_{ij(\ell)} = ((\phi\sigma)^*_{ij(\ell)} - \phi_j) + \epsilon_{ij(\ell)}.$$

Then equation (44) becomes

$$t_{ij(\ell)} = O_j + T(\Delta_{(\ell)}) + a + b\Delta_\ell + \phi_j + \epsilon^*_{ij(\ell)} \tag{45}$$

Denote an observed residual by $y_{ij(\ell)}$; that is,

$$y_{ij(\ell)} = t_{ij(\ell)} - O_j - T(\Delta_{(\ell)})$$

or

$$y_{ij(\ell)} = a + b\Delta_\ell + \phi_j + \epsilon^*_{ij(\ell)}. \tag{46}$$

Equation (46) gives a suitable model for estimating a "smooth" change in the 1968 times by employing the Veith-Clawson data.

With randomization our previous work indicates that

least squares analysis is appropriate. Since

$$\bar{y}_{\cdot j(\cdot)} = a + b\bar{\Delta}_{\cdot(\cdot,j)} + \phi_j + \bar{\epsilon}^*_{\cdot j(\cdot)}$$

where

$$\bar{y}_{\cdot j(\cdot)} = \sum_i y_{ij(\ell)}/n_j ,$$

$$\bar{\Delta}_{\cdot(\cdot,j)} = \sum_i \Delta_{\ell(i,j)}/n_j , \quad \Delta_{\ell(i,j)} = \Delta_\ell$$

and

n_j = number of observations on the j^{th} event,

we obtain the result

$$y_{ij(\ell)} - \bar{y}_{\cdot j(\cdot)} = b(\Delta_\ell - \bar{\Delta}_{\cdot(\cdot,j)}) + (\epsilon^*_{ij(\ell)} - \bar{\epsilon}^*_{\cdot j(\cdot)}). \tag{47}$$

From equation (47) we have directly that

$$\hat{b} = \frac{\sum_{i,j}(y_{ij(\ell)} - \bar{y}_{\cdot j(\cdot)})(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})}{\sum_{i,j}(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})^2}$$

$$= \frac{\sum_{i,j}(y_{ij(\ell)} - \bar{y}_{\cdot j(\cdot)})\Delta_{\ell(i,j)}}{\sum_{i,j}(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})^2} . \tag{48}$$

We are now confronted with a dilemma. It is not possible to evaluate the denominator of equation (48) from the Veith-Clawson data. Since

$$\sum_{i,j}(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})^2 \leq \sum_{i,j}(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,\cdot)})^2$$

(as can be readily shown), we have that

$$\hat{b} \geq \frac{\sum_{i,j} (y_{ij(\ell)} - \bar{y}_{\cdot j(\cdot)})\Delta_{\ell(i,j)}}{\sum_{i,j} (\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,\cdot)})^2}$$

$$= \hat{\hat{b}}, \text{ say,}$$

where

$$\bar{\Delta}_{\cdot(\cdot,\cdot)} = \frac{\sum_{i,j} \Delta_{\ell(i,j)}}{N},$$

$$N = \sum_{j=1}^{k} n_j,$$

and

k = the number of events.

We obtain the result that

$$\hat{\hat{b}} = 0.0013$$

and, therefore,

$$\hat{b} \geq 0.0013,$$

a result in direct opposition to the results of Lilwall-Douglas and Cleary-Muirhead.

Since inspection of Figure 1 indicates that the variance of $\epsilon^*_{ij(\ell)}$ decreases with increasing ℓ, employing weighted least squares would seem advisable. We employ weights proportional to the cell variances to obtain

$$\hat{b} = \frac{\sum_{\ell} \frac{1}{\sigma_\ell^2} \sum_{(i,j)} (y_{ij(\ell)} - \bar{y}_{\cdot j(\cdot)})(\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})}{\sum_{\ell} \frac{1}{\sigma_\ell^2} \sum_{(i,j)} (\Delta_{\ell(i,j)} - \bar{\Delta}_{\cdot(\cdot,j)})^2} \quad (49)$$

where
$$\sigma_\ell^2 = Var(\epsilon_{i,j(\ell)}),$$

and

(i,j) indicates a sum over all i,j at distance ℓ. It is now impossible to evaluate both the numerator and denominator of equation (49) from the Veith-Clawson data. In order to proceed further, we suppose that

$$\bar{\Delta}_{\cdot(\cdot,j)} = \frac{\sum_\ell n_\ell \Delta_\ell/\sigma_\ell^2}{\sum_\ell n_\ell/\sigma_\ell^2} = \bar{\Delta}_{\cdot}\text{, say.} \tag{50}$$

Since the σ_ℓ^2 are unknown, we shall estimate them with the data of Table 2. With these assumptions, we obtain

$$\bar{\Delta}_{\cdot} = 64.4^\circ$$

and

$$\hat{b} = -0.0032.$$

However, most of the events had stations over a wide distance range so that the assumption $\bar{\Delta}_{\cdot(\cdot,j)} = \bar{\Delta}_{\cdot}$ may not be seriously in error. Unfortunately, it is impossible to ascertain the effect of this assumption in equation (49).

Since the cell residuals are correlated, we cannot

even estimate the variances of $\hat{\beta}$ or $\hat{b}$. Thus we cannot make any tests even if the assumption of normality is made. What is needed is the Veith-Clawson data without the source terms removed. Unfortunately, these data were unavailable. In the next section we shall make an approximation that allows statistical tests to be made. For the present, qualitative conclusions can be made. There is almost certainly a decrease in variance with increasing distance. Thus more precise estimators are obtained by employing weighted least squares. A difficulty in applying weighted least squares is that the weights are not known. It is not clear how using estimated weights affects the results. Allowing for the approximations employed the estimate of the slope is 0.0013 without weights and -0.0032 with weights. If it were not for correlation the estimated weights would be acceptable since the fewest observations in a cell was 50. In view of these considerations we believe that the slope is small and negative. Further, the estimate -0.0032 is probably significant or just significant in view of the results of the cell t-tests. Note also that if -0.0032 is a significant difference then the Lilwall-Douglas value of -0.006 is significantly different from the value indicated by the Veith-Clawson data and the

value of -0.023 as given by Cleary-Muirhead is unquestionably significant, that is, on a statistical basis it is impossible to accept the value -0.023 as compared to zero. We conclude that there is possibly a slight negatively tilted correction to the 1968 P times with respect to the population represented by the Veith-Clawson data. However, we do not believe that this correction is even as large in magnitude as Douglas (1970) suggests. We shall return to this point in the next section.

Since the Cleary-Muirhead (1969) results are based on Longshot residuals it is now apparent that if the 1968 times were altered to fit the Longshot data, then they could not possibly fit the Veith-Clawson data. We emphasize the fact that how well a given world-wide average travel time function locates a specific event is not a suitable criterion of the validity of the travel-time function. In order to locate accurately and precisely the effect of the $(\phi\sigma)^*$ term must be accounted for.

With the assumed models and randomization scheme we have two possible extreme cases: (i) a simplified model does, in fact, agree with real data and measurement errors are the only errors and (ii) there is no measurement error and all errors are due to inadequacies of a specific model. The actual situation is almost certainly between these two

extremes. If we assume case (i) and a set of stations fixed before observation, we are led to a least squares location procedure in time errors (Tucker, et al., 1968). Typically in location problems the data employed arises from those stations that actually observe the event. If this is the case, then it cannot be claimed that the stations were fixed and chosen before observation. However, our results indicate that with an adequate total sample size the ordinary estimators may still be employed. On this basis, the location procedure (Tucker, et al., 1968) employed in the 1968 studies can be justified.

The Veith-Clawson data indicate that for the larger shots and earthquakes that case (ii) may be more appropriate. If we assume case (ii) we are led to a least squares location procedure in distance errors. If we consider the distance range 25° - 95° in Figure 1, we observe that the smoothed, standard deviation is very nearly proportional to the travel time derivative curve. Thus, if we divide by the travel time derivative, then we can essentially obtain a constant variance for the error term; that is, we perform a variance stabilizing transformation. Transformations of this type are frequently employed and can lead to improved results (Kempthorne, 1952).

The fact that the standard deviation is proportional to the travel time derivative is not unreasonable. Most $(\phi\sigma)^*$ terms should be small. If this is the case, then $(\phi\sigma)^*$ can be reflected in a small distance change. Longshot is possibly an extreme case and the distance error is about 25 km. A small distance change acts like a perturbation and produces a time change at a given distance proportional to the derivative. If we take a collection of these perturbed time terms then their standard deviation will also be proportional to the time derivative.

Further consideration of the standard location equations (Tucker, et al, 1968) shows that dividing by the time derivative essentially produces a location procedure in distance. For simplicity we assume that depth is known and only desire an epicenter determination. Let $\Delta(t)=T^{-1}(\Delta)$, Δ an observed distance and Δ_0 an initial distance. If (Tucker, et al, 1968)

$$\delta\tilde{\phi}=(\sin\theta)\delta\phi$$

$$\delta\tilde{\theta}=\delta\theta$$

μ_0 = the true origin time

ϕ = east longitude of the event

and

θ = seismological co-latitude of the event,

then one readily obtains the first order result that

$$\Delta_o - \Delta = \frac{\partial \Delta}{\partial t} \delta \mu_o + (-\sin Z)\delta \tilde{\phi} + (\cos Z)\delta \tilde{\theta} + e, \qquad (51)$$

where

e = the error term in distance

and

Z = the azimuth angle from the meridian through the epicenter, measured north through east, to the arc from epicenter to the station.

Equation (51) is analogous to the time equation of Tucker, et al (1968).

Since $T(\Delta)$ is monotone increasing and possesses a derivative for every distance, we have that

$$\frac{\partial \Delta}{\partial t} = 1 \Big/ \frac{\partial T}{\partial \Delta}$$

and

$$\frac{\Delta_o - \Delta}{\partial \Delta / \partial t} \doteq t - \mu_o - T(\Delta_o).$$

Thus, if we locate by iteration, depth is known and a common set of stations is employed, then location by distance is essentially equivalent to location by time.

Therefore, if the distance procedure had been employed in the 1968 studies, essentially the same estimates would have been obtained. Slight differences in location would probably have occurred since truncation was employed in the 1968 studies.

However, frequently depth is not known and must be estimated and in routine location procedures truncation must always be employed so that these two techniques, in general, yield different locations for identical initial data sets. Inspection of the Veith-Clawson data indicates that for the shorter distances it is possible to truncate a station due to model inadequacies as well as time contamination. Veith (personal communication, 1970) reports that frequently in routine locations the closer stations are, in fact, truncated. Use of the distance location procedure has improved many of these locations (Veith, personal communication, 1970). Therefore, it appears that the distance procedure may be superior to the time procedure in large scale routine location techniques.

Regardless of the final estimation scheme employed, the presence of non-normal errors (gross timing errors, etc.) must somehow be taken into account. Thus, we truncate any data set using essentially the sample median and

the mean deviation about the median. We refer the reader to Herrin, et al (1968), page 1282 for the details. The important point to note is that any truncation procedure can truncate due to contamination and also due to model misfit. In view of the data set employed in the 1968 studies we believe that the truncation due to model misfit comprised a small portion of the truncated observations. Since the distance location procedure is essentially a variance stabilizing technique, it should reduce even further the proportion of truncation due to model misfit.

ESTIMATION OF WORLD-WIDE TRAVEL TIMES AND OF FIRST ORDER STATION INTERACTION FUNCTION

Assuming that the only errors are measurement and those due to randomization over a source effect and that a set of stations was fixed before observation (Tucker, et al, 1968), we have estimated corrections to the Jeffreys-Bullen travel times (Herrin, et al, 1968) and a first order station interaction function (Herrin and Taggart, 1968a). Since not all stations see all events, we believe that the stations should be considered as having been selected at random from a fixed total set. Regardless of which randomization scheme is assumed, our present results indicate that the estimates are essentially unbiased and optimal if we assume that the events have known locations. As previously observed, if we assume normality, then the usual tests can be made.

Since most of the data in the 1968 studies were from earthquakes, the model employed must include location terms. Now the first order azimuthal source terms are (equation (21) and (22)):

$$a_j \sin Z_{ji} \text{ and } b_j \cos Z_{ji}$$

and the first order epicenter location terms are (Tucker, et al, 1968):

$$\left(\frac{\partial T}{\partial \Delta_{ij}}\right)(-\sin Z_{ji}) \text{ and } \left(\frac{\partial T}{\partial \Delta_{ij}}\right)(\cos Z_{ji})$$

so that to a first order the location and source effects are confounded. In fact, if all stations occur at the same distance from an event, then the terms are proportional and the matrix of normal equations (equations of condition) is singular. In this case, the effects are completely confounded and no estimate of each separately can be made.

In the 1968 studies, all events employed were located (even shots whose locations were known) so that there may be bias due to mislocation in the estimates obtained. Our previous analysis of the Veith-Clawson data indicates that for the travel-time estimates this is probably not the case. We now continue our investigation of possible bias in the travel-time estimates. We shall ignore the source mean removal in the Veith-Clawson data. The model becomes

$$y_{ij(\ell)} = \beta_0 + \beta\Delta_\ell + \epsilon_{ij(\ell)}$$

or

$$y_{\ell k} = \beta_0 + \beta\Delta_\ell + \epsilon_{\ell k} \qquad (52)$$

where we assume that the $\epsilon_{\ell k}$ are independent and for test purposes normally distributed. Also, equation (52) is in the form of standard regression analysis and the (weighted) least squares estimates are

$$\hat{\beta}_0 = 0.2255$$

and

$$\hat{\beta} = -0.0032.$$

Thus, $\hat{\beta}$ is equivalent to $\hat{b}$, a fact that is obvious if we consider the least squares equations. It is now possible to estimate $\sigma^2_{\hat{\beta}}$, the variance of $\hat{\beta}$. We obtain

$$\hat{\sigma}^2_{\hat{\beta}} = 1.9996 \times 10^{-6}$$

and

$$\hat{\sigma}_{\hat{\beta}} = 1.414 \times 10^{-3}.$$

Here we assume the weights are known and are equal to the estimated cell variances. The observed z is

$$z = -0.0032/0.001414 = -2.24.$$

If we choose $\alpha = .01$ the critical z is 2.576 (for a two-sided test) and for $\alpha = .05$ (two-sided) the critical z is 1.960. Thus with an α of .01 we do not reject the hypothesis that β is zero and with an α of .05 we do reject. It appears that there may be a slight correction needed to the 1968 times with respect to the population represented by the Veith-Clawson data or

$$T^* = T_{1968} - 0.0032\,\Delta$$

may be a more accurate estimate of a teleseismic shot world-wide average travel-time function. However, the results are not that conclusive and we believe that, based on the Veith-Clawson data, there is no compelling reason to alter the 1968 times. Note also that if we test the hypothesis that β is zero versus the alternative that $\beta = -0.006$ (the L-D value) we obtain analogous results. The observed z is again -2.24 but now the test is one-sided so that the critical z's are -2.326 ($\alpha = .01$) and -1.645 ($\alpha = .05$).

For comparison, we shall test

H_O: $\beta = -0.006$ (L-D)

versus

H_A: $\beta > -0.006$

and

H : $\beta = -0.023$ (C-M)

versus

H_A: $\beta > -0.023$.

For the first test the observed z is 1.98. The critical z's are 2.236 ($\alpha = .01$) and 1.645 ($\alpha = .05$). (The test is in the positive direction.) We again reach the same conclusion with respect to the value suggested by Lilwall

and Douglas (Douglas, 1970). However, for the second test the observed z is 14.00 which is significant with any reasonable α level. The value of -0.023 suggested by Muirhead and Cleary (1969) is just not consistent with the Veith-Clawson data.

To further test for possible bias due to mislocation in the travel time estimates we made a "Monte Carlo" study of island arc events in the Pacific and Atlantic. The actual events simulated and the procedure is based on a model of Longshot which will be discussed in detail in the section on the source interaction function. The event regions and events for this study are given in Table 3. We employed 20 events and 52 stations. The stations are given in Table 4. All stations employed are from the 1968 studies (Herrin and Taggart, 1968a). We attempted to make the station selection both world-wide and representative of those stations for which we have estimated station corrections. The event regions are basically island arc regions and all but two are in the Pacific. We attempted to make these regions representative of the data set employed in the 1968 studies. Most of the event regions are in the western Pacific which is consistent with the 1968 data set (Herrin, et al., 1968).

The Puerto Rican event is represnetative of only one event near the Lesser Antilles and as such might not have been included. To the extent that these events and stations are representative of the 1968 data the results will indicate the possible effects of mislocation in travel time bias. We tacitly assumed that event locations outside of island arc regions in the 1968 studies were unbiased and hence no data for these regions is included in the study. To give an extreme case we assumed that a source function equivalent to Longshot existed in each event region. This source function approximates that of the Longshot model study and is given by $[-0.5 + 1.32 \sin (Z_{ji} + 273.6^{\circ} + \theta_k)]$ (Herrin and Taggart, 1968b) where θ_k gives an adjustment to Z_{ji} for the trend of the island arc. (For the Longshot event $\theta_k = 0.$) With this source function and the 1968 times we generated observed arrival times at each of the stations. No additional random errors were introduced so that the study is Monte Carlo in nature only to the extent that the event regions and stations constitute a random sample from the 1968 data set.

The generated arrival times were then employed in locating the events. Time residuals were obtained from these

estimated locations. The residuals were then grouped on a 5° cell basis to compute the mean and standard deviation and the results are presented in Table 5. The means given in Table 5 can be thought of as estimated corrections to the 1968 times. We thus take as true the 1968 times and use data that has as error only model misfit or source interaction. We then simulated the first iteration in a procedure to estimate the true travel times. This should represent an extreme case since only one step in the iteration was performed. In finding the final estimated curve we would apploy the corrections given in Table 5 to obtain a new travel-time curve, then re-estimate the locations, etc. The data of Table 5 represent that portion of estimated corrections that would be due to mislocation and as such are over-corrections. Any estimates that have this phenomenon included should have the corrections of Table 5 subtracted in order to eliminate mislocation bias. The results of Table 5 indicate that the estimated curve would be slow with respect to the true curve from 20° to 60°, unchanged in the range 60°-75° and fast in the range 75°-100°. The estimated slope using weighted least squares with the observed variances as weights yields a slope estimate of

-0.0030. However, this is a positive slope with respect to the Veith-Clawson, Lilwall-Douglas or Cleary-Muirhead data and indicates that any rotation produced in the 1968 times from this effect is in a direction opposite to that suggested by Douglas (1970) and Muirhead and Cleary (1969). While only one iteration was made, it is unlikely that further iterations would change the slope from +0.0030 to -0.0032 or more negative values. If the Longshot model represents an extreme case of source function and the selected events are representative of the 1968 study then we conclude that any bias due to mis-location should be small. The small negative slope indicated by the Veith-Clawson data could be due to effects other than mislocation.

As a final check we can compare the 1968 times to the Lilwall and Douglas (1970) times. Employing the smoothed times and 95% confidence intervals as given by Lilwall and Douglas (1970) for their estimated travel-time curve and the 1968 times and 95% confidence intervals (Herrin, et al., 1968), one can perform a significance test. Ignoring a baseline shift the confidence bands of the two curves (Lilwall and Douglas (1970) curve considered at only their smoothed points) fail to touch or overlap at only the distance 43°.

The separation is approximately 0.1 sec. Since adjacent confidence intervals are not independent, it is impossible to determine the significance level of the test. Further, the data sets overlap somewhat, and many criteria present themselves. We could reject the hypothesis that the two curves represented by the two data sets are equivalent if one or more confidence intervals do not touch or overlap or if two or more do not touch or overlap, etc. It is not clear which criterion is appropriate. We submit that the data do not suggest that the two curves differ seriously.

Lilwall and Douglas (1970) and Douglas (1970) claim that the rotation between the L-D times and the 1968 times is possibly due to event mislocation. In support of this claim, they simulated one iteration of the 1968 studies with their data. However, it is not correct to compare on just one iteration. One iteration gives estimates of the travel times that are unadjusted for location and other model effects. Any unbiased estimates must be adjusted for all effects. This is analogous to attempting to invert a matrix and not accounting for the off diagonal terms. The total estimation scheme employed in the 1968 studies (Tucker, et al., 1968) was essentially a Siedel iterative procedure (Householder,

1953). The Siedel procedure will converge; that is, give the same estimates as matrix inversion if certain conditions are met. We believe these conditions were met in the 1968 studies and probably for the Lilwall-Douglas (1970) data. If this is the case, then the joint epicenter method of Lilwall and Douglas and the Siedel method of the 1968 studies are mathematically equivalent. In order to compare correctly the two methods, the Lilwall-Douglas data should have been iterated enough times to be equivalent to the twelve iterations performed in the 1968 studies.

We conclude that there is no appreciable bias in the 1968 times and, hence, at present no compelling reason to alter these times. As more shot data or data from large, well located earthquakes becomes available, it should be possible to estimate any change in the 1968 times. Since it appears that a slight rotation may be needed to make the 1968 times more nearly fit a teleseismic shot average, we offer a possible explanation. The indicated change is probably due to many sources. We first believe that there may be real differences between a travel-time curve for the population represented by the Veith-Clawson shot data and the population represented by the 1968 data. Inspection of

Table 1 indicates that only Longshot or approximately 5% of the data come from active seismic regions. Thus the Veith-Clawson population is primarily from aseismic areas while that of th 1968 studies is primarily from earthquake zones. The important result, we believe, is that the Veith-Clawson data indicate that the 1968 times, based essentially on earthquake data, are also applicable to shots in aseismic regions. Also there is possibly a slight rotation due to mislocation. However, part of any rotational change may be due to the arbitrary model employed for the distance range 0°-20° in the 1968 studies. Inspection of the distance ranges 10°-15° and 15°-20° of the Veith-Clawson data indicates that the 1968 times are highly significant and fast. Since a change in the model could produce a slight rotation (Herrin, et al., 1968) and the effects of the arbitrary model may go out to 30°, a different initial model can account for part of any rotational change. Finally, regional variations in P times may occur at distances as great as 60°. If this is the case, then estimates in the range 25°-60° could vary widely depending on the actual data set.

While there is apparently only slight or no bias in the 1968 times, the results on station terms are not as conclusive.

With the exception of the Lilwall-Douglas (1970) work, the station corrections show the same overall features as other published corrections. The Lilwall-Douglas corrections for European stations appear to have about a 0.5 sec. negative displacement with respect to those of the 1968 studies. This shift may or may not be statistically significant. Further, the randomization analysis indicates that a minimum of 50 to 100 observations per station are required in order to have the order of the bias be 0.02 sec or less. We believe that those station terms of the 1968 studies based on 50 or more observations are adequate. We reason as follows: To the extent that there is no travel time bias, the location terms were completly confounded with any first order azimuthal source terms and hence were actually estimated. We cannot separate a mislocation and an azimuthal source effect. Thus, any estimable model will not have separate location and azimuthal source terms. One or the other or some linear combination of the two will suffice. Since a mislocation accounts for nearly all of a first order azimuthal source effect, there are no bias terms in the model; that is, to a first order a possible adequate model has (i) location terms, (ii) station effect terms and (iii) possibly travel time terms.

PROBLEM OF THE SOURCE INTERACTION FUNCTION

Since the location terms are confounded with the first order azimuthal terms of a source interaction function, in general, we mislocated in the presence of strong source terms. While this was desirable for the 1968 studies it is undesirable for a pure location problem. In order to investigate the possibility of estimating a source function and a location simultaneously we performed a Monte Carlo study using the estimated Longshot source function (Herrin and Taggart, 1968b) and a network of stations from those employed in the 1968 studies. The network was balanced to the extent possible and had stations well distributed in azimuth and distance from the source. We assumed an event located on Amchitka. No station terms were included or estimated. The 46 stations employed are given in Table 6.

The study was performed as follows: With the fixed network the Longshot model was used to derive time terms for each station. These were added to the true travel times (as given by the 1968 P times) from event to station. This constituted a fixed set of initial arrival times for each station. To these times were added independent normal deviations with constant variance σ^2,

for all stations. The values of σ =0, 0.08, 0.8 and 1.0 sec were employed. The initial arrival times plus the normal error terms were employed as a set of observed arrival times at each station. These observed arrival times were employed in a joint location and source term estimation procedure. The actual location was recorded and a series of runs made for each variance level. In the case of σ =0 only one run was made. The location was 0.065 kms north and 0.005 kms west of the true location, that is, with no error we can locate and estimate a source function simultaneously. The matrix of normal equations is not singular in a neighborhood of the true location. For the other variance levels repeated runs were made. The mean location change and standard deviation (both in kms) were obtained for the north-south and east-west directions about the true location. The results are presented in Table 7 where

$$\tilde{\sigma}_{N-S} = \sigma_{N-S}/\sigma \text{ and } \tilde{\sigma}_{E-W} = \sigma_{E-W}/\sigma .$$

To check the validity of the results certain tests were made. Since multiple minima may occur in nonlinear least squares an extensive search was made of a selected run. For this run the location without source terms was 16.64 km north

and 9.29 km east while the location with source terms was 29.81 km south and 5.50 km west. The location without source terms is apparently typical while the location with source terms appears to be too far south. If spurious minina are commonly reached this run should serve as an example. The search indicated that the iterative least squares procedure did in fact reach the true minimum and that there there is only one minimum. This is not unexpected since the travel time nonlinearities are not great. Throughout the study different initial points were employed on the same run in order to check for dependence on the initial point. In all cases the final locations were identical. The more common initial point was the true location since this minimized computer run time.

The last set of runs given in Table 7 also serve as a check. For this group no source function was estimated. This results in locations biased toward the north. The "true" location, that is, the location without error is given by the run with $\sigma = 0$ and is 22.46 km north and 1.96 km west. The mean shifts of -0.715 and -0.179 are not statistically significant with

$\alpha = 0.05$ or $\alpha = 0.01$. In fact the N-S shift is barely significant at the 10% level and the E-W shift is not significant until the 63% level is reached. Here we test each direction independently. This is approximately valid since the network of stations is nearly balanced. The off diagonal terms in X'X are smaller by an order of magnitude than the diagonal terms and the confidence ellipse is essentially circular. Practically, the north-south and east-west locations are statistically independent. Based on this consideration the results were presented for the individual axes in Table 7.

Since these checks indicate that the Monte Carlo runs were properly made we shall now investigate the implications of the results. First we see that the nonlinearities produce no apparent effect on the variance of the estimated locations. All of the $\hat{\sigma}$ lie essentially between 15 and 20 km for those runs where source terms were estimated. Also the standard deviations for the non-source runs were consistent with real data for large events located by good stations with balanced networks. Standard deviations of 5-7 km are commonly observed. It appears that introducing source terms into the given network

produces a fourfold increase in the standard deviation of the estimated location. This is not unexpected since even for a well balanced network the location terms and azimuthal source terms are partially confounded. The normal equations while apparently not singular have near-singular tendencies.

Secondly we note that the estimated location is affected by model nonlinearities. As the error standard deviation is increased the bias increases. This is not uncommon in non-linear regression. Investigation of the $\sigma = 0.08$ case indicates that the bias is not significant for any reasonable α. In fact it appears that for balanced networks if the only errors were observational, then we could simultaneously estimate a location and source effect.

Finally, with the balanced network studied it appears infeasible to estimate both a location and azimuthal source effect. The $\sigma = 0$ case yields a bias of 22.46 km and a standard deviation of 3.99 km for the north-south direction and a bias and standard deviation of 1.96 km and 3.43 km respectively for the east-west direction. The corresponding mean square errors are 520.36 km^2 and 15.61 km^2 for the respective directions or

an overall MSE of 535.97 km^2 based on independence. If we take σ = 1.0 for a typical location, then similar computations yield for the source estimated case 339.59 km^2 and 294.61 km^2 MSE for the respective directions for an overall MSE of 634.20 km^2. Thus the source estimated case does not compare unfavorably. However, let us now suppose that there is no true source effect present. As is indicated by the last case in Table 7 and by the work of Tucker, et al., (1968) in the source-not-estimated case the bias would be zero and the overall MSE 27.56 km^2. However, when the azimuthal source is estimated the procedure no longer acts linear. It is not clear what the bias would be. A conservative estimate is zero. Further the variances should remain the same since the nonlinearitie apparently do not affect them. With these assumptions the MSE for the source estimated case with no source function present is 633.89 km^2. In this case it would be unreasonable to estimate simultaneously a location and an azimuthal source effec

These results are based on an excellent network. Practical networks are often poorly balanced. It is of interest to compare the two procedures under these conditions. To

simplify matters and reduce run time only the zero error case was studied. The networks designated by I, II, III and IV and given in Table 8 were studied. The results are presented in Table 9. As one might expect the source estimation procedure worked reasonably well without observation error (σ=0). However, for network IV the procedure yielded a N-S error of 72.93 km as compared to 16.61 km when locating without source terms. This is clearly unacceptable. Further, network IV is similar to network III. In fact IV was selected while attempting to obtain a network that was a reflection of III. There is no indication in the two networks that the N-S error would go from 0.45 km to 72.93 km. There is no way in practice to know that this type of phenomenon has occurred.

We pursue these considerations by now considering a real event, Longshot, which was large and well recorded. The Longshot observations were initially sorted on the estimated station variances and the number of observations per station as given by Herrin, et al., (1968). The criteria were: minimum number of observations 25 and maximum variance 1.00 sec^2. From this sorting two reasonably balanced networks were

selected. They are given in Table 10 and are labled V and VI. The fewest observations for V was 34 and the largest variance was 1.00 sec^2. For VI these quantities are 33 and .92 sec^2 respectively. With these networks a number of combinations of location procedures were investigated. For each network we located with and without station corrections (Herrin, et al., 1968) and with and without estimates of azimuthal source corrections. The results are presented in terms of north-south and east-west errors in km in Table 11 and are discouraging. The station corrections appear to work reasonably well when a source function is not estimated. Unfortunately, when we estimate an azimuthal source function the station corrections appear to further degrade the location. However, we note that generally the corrections are in the right directions, even when all are applied or estimated; that is, the corrections move the estimated location south. Network V gives a location 19.86 km north of the true and VI gives a location 30.22 km north of the true. The apparent true location is about 25 km north of the true. Errors (both measurement and model) in the data for these two networks are apparently such that "precise" solutions

are obtained without any corrections and "imprecise" solutions are obtained with corrections. That is, the Longshot source function is hidden in the error, a result consistent with the Monte Carlo studies. We conclude that there appears to be no practical way of directly estimating an azimuthal source function and a location simultaneiously unless the station error variances are small and any corrections (station corrections) that are applied are accurate and precise.

Since our direct attempts to both locate and estimate an azimuthal source function have proved to be unsuccesful we have investigated certain indirect methods. A promising approach involves the use of deterministic source models for the regions in which source bias is expected to occur, namely the island arcs. The model we have used for the Amchitka region (Herrin and Sorrells, 1969) is that of a down-going high velocity plate of upper mantle material. This model, consistent with the concepts of ocean-floor spreading and global plate-tectonics, is represented in profile by Figure 3. The dip angle and length of the plate was determined by the spatial distribution

of earthquakes in the central Aleutian Islands. Using a three-dimensional ray tracing program developed by G. Sorrells, we calculated the travel time residual pattern relative to the 1968 travel times which would result from this distribution of velocities beneath Amchitka. Figure 4 shows the residuals as contours; the map is symmetrical about the north-south line so that only the eastern half is plotted. Note the complexity of the map and that the most negative residuals, representing the fastest ray paths, are at distances of 20° to 60° and azimuths northeast and northwest from Longshot. In Figure 5 the residuals are plotted as a function of distance for four azimuths. Using the actual locations of 84 stations which recorded Longshot we computed the station residuals which would result directly from this model. Figure 6, which shows these residuals plotted as a function of azimuth, is very similar to the same kind of plot of observed Longshot residuals. A first order sine-curve was fit to the model residuals of the form $A + B\sin(Z_{ji} + \epsilon)$ where ϵ is the phase angle.

The resulting values were $[-0.32 + 0.68 \sin(Z_{ji} + 250°)]$ which can be compared to the fit to observed Longshot residuals which was, $[-0.5 + 1.32 \sin(Z_{ji} + 273.6°)]$. These functions are similar; however, the observed data gave larger values for both the A term and the amplitude of the sine function. The similarity can be improved by increasing the velocity contrast in the assumed source model or by lengthening the down dipping plate.

The source parameters $\sigma_{j\ell}$, $A_{j\ell}$ and $B_{j\ell}$ in the first order model were calculated for the source region represented by Figure 3. We then estimated the Longshot epicenter using equation (20) as a statistical model, with the distance dependence of the station terms, $\phi_{i\ell}$, $A_{i\ell}$ and $B_{i\ell}$ being neglected. The resulting location shown on Figure 7 as a dot marked "corrected times" is closer to the true location than our previous estimate. Had the length of the dipping plate in our assumed source model been 300 km instead of 200 km, the computed location would have been very close to the true shot point.

The use of deterministic source models appears to be the most promising course to follow in attempting to remove bias from epicentral estimates employed in the 1968 studies. Our previous estimates of azimuthally dependent station corrections may well be biased because of the spatial correlation of source effects along the island arcs of the Pacific and Atlantic. The next step in a re-estimation procedure is to relocate all of the island arc events we have used, assuming the travel time tables are correct, based on a modified equation (20) and assumed source models for each arc. (The source terms are modeled rather than estimated.) New residuals then obtained can be used to re-estimate the station terms. The process must involve serveral interations because the time anomaly field for a given source model is very sensitive to a shift in location normal to the strike of the dipping plate.

CURRENT CAPABILITIES IN THE ESTIMATION OF LOCATION PARAMETERS

The Monte Carlo study of Amchitka, the Longshot study analysis of the Veith-Clawson data and the work of Tucker, et al., (1968) indicates that an error variance of 1 sec^2 is not unreasonable for present network capabilities and larger events. This reflects, primarily, errors due to

(i) higher order model inadequacies and

(ii) basic station operational capabilities (timing, reading, etc.).

The overall value of 1 sec^2 is only justified on a suitable randomization basis since model inadequacies are not true random variables. Further, in order to achieve this value we must routinely truncate our data sets to eliminate contamination errors. Since the first error source, (i), is approximately an order of magnitude larger than the second source, (ii), (for the better stations) any reduction in σ^2 must come from more accurate models.

As a first step the distance dependence in station correction terms should be accounted for. The Amchitka Monte Carlo study indicates that if more accurate and precise estimates of the station time terms were available it might be possible to simultaneously locate and estimate azimuthal

source terms. That is, if we can get the magnitude of error source (i) comparable to that of source (ii) we can possibly eliminate by estimation the bias due to an azimuthal source function. However, if an approximate, deterministic source model is available (estimated or known from other work), then the value of 1 sec^2 is operationally adequate as evidenced by the work of Tucker, et al., (1968), Herrin and Taggart (1968b) and the model studies of our present work. All of these studies indicate that for a well distributed network composed of 25 or more good stations, the major contributor to location error is source bias produced by an azimuthal source function. In the Herrin-Taggart (1968b)study the improvement in locations produced by use of station corrections was of the order of a few km while the improvement by use of source corrections was of the order of tens of km. The Monte Carlo studies of Tucker, et al., (1968) gave similar results. Our present model studies indicate that this phenomenon is due to the confounding of the location terms and azimuthal source terms employed in the least squares estimation procedure. Thus any real improvement in location capabilities must come from adequately accounting for azimuthal source terms.

We have as direct implications of the Monte Carlo study

of Amchitka and the Longshot studies that

(i) bias in a simultaneous location and azimuthal source term estimation procedure is of the order of a few km,

(ii) in locating either with or without azimuthal source terms, the procedure is linear with respect to the variance of the estimated location which implies that the location diagonal terms of $(X'X)^{-1}$ increase when estimating azimuthal source terms,

(iii) in either case a suitable measure of network effectiveness is given by $(X'X)^{-1}$, and

(iv) for a location without source term estimation, using a balanced network of 25 or more good stations, standard deviations of 5 to 7 km can be obtained.

Two possible solutions to the source problem are now apparent. In one we can use $(X'X)^{-1}$ as a figure of merit and attempt to minimize the overall MSE by estimating only a proportion of any azimuthal source terms. On the other hand we can model the local structure, e.g., Longshot, and use this model to predict residuals. Both have advantages and disadvantages. The first requires no specific model. It is, however, limited to a few terms in an expansion and has the drawback of estimating only a proportion of any possible

lower order azimuthal source terms. This is partially compensated by the fact that first order terms should be adequate and a measure of the overall MSE is available. The proportion of any source terms estimated is known and can be used to predict the bias and also the variance of the location estimators is known from $(X'X)^{-1}$ (assuming $\sigma^2 = 1 \text{ sec}^2$). The second method has an overall variance that is known and adequate in view of (iv). Unfortunately any bias removal is a direct function of the model and for many source regions we have no knowledge of proper models. Not only is the model important but also the location of the event with respect to the model. For example, with Longshot a change in the epicenter of about 50 km to the south should virtually eliminate all of the time residuals due to the Longshot source function. Also any location procedure would be costly since a search technique must be employed. (Analytic representations of the models we have considered are difficult to derive.)

Either of these methods can be equally effective. Our Longshot model studies indicate that a reasonable model accounted for about one-half of the bias. Assuming a standard deviation of 6 km we obtain an overall MSE of 144 km^2 + 36 km^2 or 180 km^2. To evaluate the first method let us

assume that a proportion of one-half yields a standard deviation of 10 to 12 km, say 11 km. Preliminary work with the proportion method locating the Monte Carlo Amchitka event indicates that these assumptions are not unreasonable. With these assumptions we obtain an overall MSE of 144 km^2 + 121 km^2 or 265 km^2 or slightly poorer than the model method. However, both MSE's are comparable. Thus we conclude that with proper modeling or careful proportional estimation of source terms epicenter location MSE's of 180 to 265 km^2 are obtainable.

Further, improvement in these values should be obtained by employing the distance least squares method. We have previously discussed the usefulness of the distance method in routine location procedures where truncation is employed. Recall that the distance method is essentially the result of a variance stabilizing transformation. To the extent that the distance method produces a constant variance it will, in theory, result in lower variances for the location parameters. This is so since in the case of unequal variances a weighted least squares is optimum. Thus the distance method should improve least squares depth estimates by retaining closer stations and should reduce the variance of all location parameters.

It is difficult to assess the quantitative affects of the distance method on epicenter location errors; however, the improvement in standard deviation should be only a few km. If we consider the model case and suppose that the standard deviation reduction is 2 km we obtain 144 km^2 + 16 km^2 or an overall MSE of 160 km^2. Similar considerations for the proportional case yield 144 km^2 + 81 km^2 or an overall MSE of 225 km^2. We then obtain overall epicenter location MSE's of 160 to 225 km^2. Again we see that bias reduction is the important consideration.

By employing "good" station corrections a further reduction of the location parameter variances is possible. These improvements are again slight with respect to the possible improvement in location bias. We conclude that by employing a distance location procedure, "good" station corrections and either a source model or the proportional method that for a well distributed network composed of 25 or more good stations epicenter location MSE's of less than 200 km^2 are obtainable. Furthermore, if a region can be calibrated, say, by a number of shots spaced throughout the region, then it is possible to obtain location MSE's of less than 50 km^2. In this case there is essentially no source bias and the errors are due to either measurement errors or higher order model inadequacies associated with the stations.

SUMMARY OF CONCLUSIONS

The following paragraphs provide a brief summary of the most important conclusions reached in this report.

1. It is possible to develop simplified, practical, mathematical models for the prediction of teleseismic travel times. These models lead to a non-trivial definition of a "world-wide-average" travel time function.

2. A randomization procedure for the simplified travel time models provides the basis for estimation of model and location parameters under two extreme conditions:

(a) All errors result from inaccuracies of measurement, in which case the mean square time error is minimized for a location estimate, and

(b) all errors result from model inadequacies, in which case the mean square distance error is minimized for a location estimate. In practice, errors arise from both source, and either estimation technique will yield satisfactory results. Minimization of mean square distance errors may be preferable when data from only very good stations are available.

3. When estimating travel times using combined data from explosions and earthquakes, all the sources should be treated

in the same manner. That is, the epicenters should be estimated for the explosions. Utilization of known location parameters for the explosions can, in this case, lead to bias in the travel time estimates.

4. The 1968 P-travel times (Herrin, et al., 1968) are essentially unbiased and statiscally equivalent to the most complete set of shot derived travel times (Veith and Clawson, 1972).

5. The estimated station corrections given by Herrin and Taggart (1968a) are probably adequate for those cases based on 50 or more observations.

6. Source location error is essentially confounded with the first order azimuthal source terms. This effect (source bias) may be the largest cause of error in estimating epicenters.

7. It is probably useful to re-estimate station corrections to remove any bias due to mislocation of the sources and to allow for distance dependence.

8. With "good" station corrections, methods can probably be developed for estimating epicenters with mean-square-errors of less than 200 km^2, provided the network consists of 25 or more well distributed stations. These methods will require the use of deterministic source models or the proportional method for approximating the source interaction term.

TABLE 1*

SUMMARY OF EXPLOSION DATA

Site or Event	No. of Events	No. of Observations Total	Teleseismic
Semipalatinsk	7	528	522
Algeria	6	463	442
Novaya Zemla	9	257	251
LONGSHOT	1	234	232
SHOAL	1	100	46
Nevada Test Site	6	531	313
SWORDFISH	1	48	38
CHASE II	1	36	33
CHASE III	1	53	32
CHASE IV	1	48	31
CHASE V	1	63	31
CHASE VII	1	53	33
Bikini	4	21	21
Eniwetok	3	19	19
Totals	43	2454	2044

*From Table 1, Veith and Clawson (1972).

TABLE 2

ANALYSIS OF THE EXPLOSION RESIDUALS FOR THE HERRIN TRAVEL TIMES

Mean values calculated per 5 degree block

Cell	N	Mean	SD	Variance	SD Mean	T
0-5	100	.262	2.976	8.856	.298	.880
5-10	135	.026	2.928	8.573	.252	.103
10-15	104	.955	2.924	8.551	.287	3.330
15-20	102	1.050	2.337	5.462	.231	4.537
20-25	101	.357	1.722	2.966	.171	2.086
25-30	100	.083	1.713	2.933	.171	.485
30-35	109	.367	1.388	1.927	.133	2.760
35-40	100	.392	1.237	1.529	.124	3.170
40-45	86	.112	1.208	1.459	.130	.857
45-50	72	.072	1.019	1.039	.120	.601
50-55	56	-.300	.988	.976	.132	-2.272
55-60	70	-.116	1.085	1.176	.130	-.893
60-65	84	-.144	1.212	1.470	.132	-1.089
65-70	90	-.162	1.091	1.191	.115	-1.410
70-75	80	.041	1.001	1.003	.112	.368
75-80	77	-.296	1.092	1.192	.124	-2.380
80-85	85	-.140	.935	.873	.101	-1.381
85-90	80	-.043	.948	.899	.106	-.401
90-95	68	.068	.900	.810	.109	.620
95-100	50	.364	.904	.817	.128	2.847

TABLE 3

PACIFIC-ATLANTIC MODEL STUDY EVENT REGION

Event Code	Event Region
Amchitka	Amchitka Is.: Aleutian Trench
Komandorskie	Komandorskie Is.:
Kuril	Kuril Is.: Kuril-Kamchatka Trench
Japan	Japan: Japan Trench
Ryukyu	Ryukyu Is.: Shoto Trench
Manila	Philippines: Philippine Trench
Western New Guinea	Western New Guinea
Tonga	Tonga Is.: Kermadec - Tonga Trench
Chile	Chile: Peru-Chile Trench
Lima	Peru: Peru-Chile Trench
Columbia	Columbia: Pacific Coast
Guatemala	Guatemala: Pacific Coast
Mazatlan	Mexico: Pacific Coast
Kodiak	Kodiak Is.
Unalaska	Unalaska Is.: Aleutian Trench
Mariana	Mariana Is.: Mariana Trench
Ogasawara	Japan: Japan Trench
Java	Java: Java Trench
Sandwich Is.	So. Sandwich Is.: South Sandwich Trench
Puerto Rico	Puerto Rico: Puerto Rico Trench

TABLE 4

PACIFIC-ATLANTIC MODEL STUDY STATIONS

Code	Station	Code	Station
AFI	Afiamalu, Samoa	PAL	Palisades, New York
ATH	Athens, Greece	PER	Perth, Australia
BAG	Baguio, Philippines	PVC	Port Vila, New Hebrides
BRK	Berkley, California	QUE	Quetta, W. Pakistan
BOG	Bogota (WWNSS), Columbia	REY	Reykjavik, Iceland
BOK	Bokar, India	ROM	Rome, Italy
BUL	Bulawayo, Southern Rhodesia	SAP	Sapporo, Japan
CAN	Canberra, Australia	SEM	Semipalatinsk, USSR
CTA	Charters Tower, Australia	SEN	Sendai, Japan
CMO	College, Alaska	SHI	Shiraz, Iran
CSC	Columbia, So. Carolina	SIM	Simferopol, USSR
DAL	Dallas, Texas	SOD	Sodankyla, Finland
GUA	Guam, Guam	SAN	Santiago, Chile
HLW	Helwan, Egypt	THU	Thule, Greenland
HVO	Hawaiian Volcano Obs., Hawaii	TIK	Tiksi, USSR
KIS	Kishinev, USSR	TOL	Toledo, Spain
LAN	Lanchow, China	TRN	Trinidad, West Indies
LPS	LaPalma, El Salvador	UPP	Uppsala, Sweden
LAR	Laramie, Wyoming	VAN	Vannovskaya, USSR
LHA	Lhasa, China	VCM	Vera Cruz, Mexico
LPB	La Paz, Bolivia (WWSS)	VIE	Vienna, Austria
LWI	Lwiro, Congo	VLA	Vladivostok, USSR
MED	Medan, Sumatra	WAR	Warsaw, Poland
MNT	Montreal, Canada	WEL	Wellington, New Zealand
MOS	Moscow, USSR	WIL	Wilkes, Antarctica
MBC	Mould Bay, Canada	WIN	Windhoek, South Africa

TABLE 5

PACIFIC-ATLANTIC MODEL STUDY

Cell	Mean	SD Mean	N
20-25	.035	.449	26
25-30	.173	.314	20
30-35	.185	.227	25
35-40	.041	.193	29
40-45	.129	.162	23
45-50	.018	.147	32
50-55	.016	.106	38
55-60	.040	.078	39
60-65	.008	.063	40
65-70	.006	.066	49
70-75	.006	.119	53
75-80	-.021	.172	51
80-85	-.049	.177	56
85-90	-.121	.205	57
90-95	-.073	.256	44
95-100	-.060	.256	51
100-105	.004	.259	51

TABLE 6

STATIONS EMPLOYED IN THE AMCHITKA MONTE CARLO STUDY

Code	Station	Code	Station
AFI	Afiamalu, Samoa	PAL	Palisades, New York
ATH	Athens, Greece	PER	Perth, Australia
BAG	Baguio, Philippines	PVC	Port Vila, New Hebrides
BRK	Berkely, California	QUE	Quetta, West Pakistan
BOG	Bogota (WWNSS), Columbia	REY	Reykjavik, Iceland
BOK	Bokar, India	ROM	Rome, Italy
CAN	Canberra, Australia	SAP	Sapporo, Japan
CTA	Charters Tower, Australia	SEM	Semipalatinsk, USSR
CMO	College, Alaska	SEN	Sendai, Japan
CSC	Columbia, South Carolina	SHI	Shiraz, Iran
DAL	Dallas, Texas	SIM	Simferopol, USSR
GUA	Guam, Guam	SOD	Sodankyla, Finland
HLW	Helwan, Egypt	THU	Thule, Greenland
HVO	Hawaiian Volcano Obs., Hawaii	TIK	Tiksi, USSR
KIS	Kishinev, USSR	TOL	Toledo, Spain
LAN	Lanchow, China	TRN	Trinidad, West Indies
LPS	LaPalma, El Salvador	UPP	Uppsala, Sweden
LAR	Laramie, Wyoming	VAN	Vannovskaya, USSR
LHA	Lhasa, China	VCM	Vera Cruz, Mexico
MED	Medan, Sumatra	VIE	Vienna, Austria
MNT	Montreal, Canada	VLA	Vladivostok, USSR
MOS	Moscow, USSR	WAR	Warsaw, Poland
MBC	Mould Bay, Canada	WEL	Wellington, New Zealand

TABLE 7

AMCHITKA MONTE CARLO STUDY RESULTS

n	σ	N-S Mean	σ_{N-S}	$\tilde{\sigma}_{N-S}$	E-W Mean	σ_{E-W}	$\tilde{\sigma}_{E-W}$
			Source Estimated*				
128	1.0	-4.261	17.93	17.93	-0.560	17.16	17.16
48	0.8	-3.876	12.11	15.14	1.810	16.28	20.35
48	0.08	0.126	1.269	15.86	-0.015	1.763	16.60
			Source Not Estimated*				
88	1.0	-0.715	3.988	3.988	-0.179	3.431	3.43.

*North and east are plus while south and west are minus.

TABLE 8

TEST NETWORKS FOR AMCHITKA MONTE CARLO STUDY

Network	Station
I	Berkeley, California Hawaiian Volcano Obs., Hawaii Laramie, Wyoming Mould Bay, Canada Sapporo, Japan Sendai, Japan Thule, Greenland Tiksi, USSR Vladivastok, USSR
II	Afiamaul, Samoa Columbia, South Carolina Lhasa, China Moscow, USSR Palisades, New York Port Vila, New Hebrides Uppsala, Sweden Vera Cruz, Mexico
III	Berkeley, California Laramie, Wymoming Mould Bay, Canada Thule, Greenland Tiksi, USSR
IV	Hawaiian Volcano Obs., Hawaii Sapporo, Japan Sendai, Japan Tiksi, USSR Vladivastok, USSR

TABLE 9

AMCHITKA TEST NETWORK ERRORS

Network	No. of Stations	Source not Estimated*	Source Estimated*
I	9	17.21 -0.95	-0.15 0.05
II	8	23.15 -1.34	-1.31 0.74
III	5	18.05 -0.06	0.45 -0.06
IV	5	16.61 -1.43	72.93 -3.11

*North and east are plus while south and west are minus.

TABLE 10

LONGSHOT STUDY NETWORKS

Network	Code	Station	Code	Station
V	ALQ	Albuquerque, New Mex.	LAR	Laramie, Wyoming
	BNS	Bensberg, Germany	MAT	Matsushiro, Japan
	CAN	Canberra, Australia	MNT	Montreal, Canada
	CTA	Charters Tower, Australia	NUR	Nurmijarvi, Finland
	CMC	Copper Mine, Canada	PAS	Pasadena, Calif.
	COL	College Outpost, Alaska	PMG	Pt. Moresby, New Guinea
	CPO	Cumberland Plateau S.O., Tennessee	PRU	Pruhonice, Czech.
	DAL	Dallas, Texas	QUE	Quetta, W. Pakistan
	DH-	Delhi, New York	RES	Resolute Bay, Cana.
	FBC	Frobisher, Canada	SHI	Shiraz, Iran
	ISO	Isola, France	SOD	Sodankyla, Finland
	JER	Jerusalem, Israel	STR	Strasbourg, France
	KJN	Kajaani, Finland	TOO	Toolangi, Australia
	KHC	Kasperske Hory, Czechoslovakia	TRO	Tromsoe, Norway
	KEV	Kevo, Finlan	UPP	Uppsala, Sweden
	KIR	Kiruna, Sweden	VAL	Valentia, Ireland
	KON	Kongsberg, Norway	WMO	Wichita Mts. Seis. Obs., Oklahoma
VI	BKS	Byerly, California	LON	Longmire, Wash.
	BUT	Butte, Montana	MAT	Matsushiro, Japan
	COL	College Outpost, Alaska	MN-	Mina, Vevada
	EUR	Eureka, Nevada	PRS	Paraiso, Calif.
	FBC	Frobisher, Canada	PAS	Pasadena, Calif.
	GOL	Golden, Colorado	SOD	Sodankyla, Finland
	KEV	Kevo, Finland	SPO	Spokane, Washington
	KIR	Kiruna, Sweden	TFO	Tonto Forest Seis. Obs., Arizona
	LAR	Laramie, Wyoming	TUC	Tucson, Arizona
	LAW	Lawrence, Kansas	UBO	Uinta Basin Seis. Obs., Utah
	LC-	Las Cruces, New Mex.	VIC	Victoria, Canada
	MHC	Lick (Mt. Hamilton), California	WMO	Wichita Mts. Seis. Obs., Oklahoma

TABLE 11

LONGSHOT STUDY NETWORK ERRORS

Network			North-South Error*	East-West Error*
V	Source Not Estimated	No Station Corrections	19.86	-1.66
		Station Corrections	21.08	-1.50
	Source Estimated	No Station Corrections	-19.74	1.28
		Station Corrections	-35.51	10.19
VI	Source Not Estimated	No Station Corrections	30.22	-2.57
		Station Corrections	21.03	2.05
	Source Estimated	No Station Corrections	-9.40	-53.11
		Station Corrections	-66.28	6.72

*North and east are plus while south and west are minus.

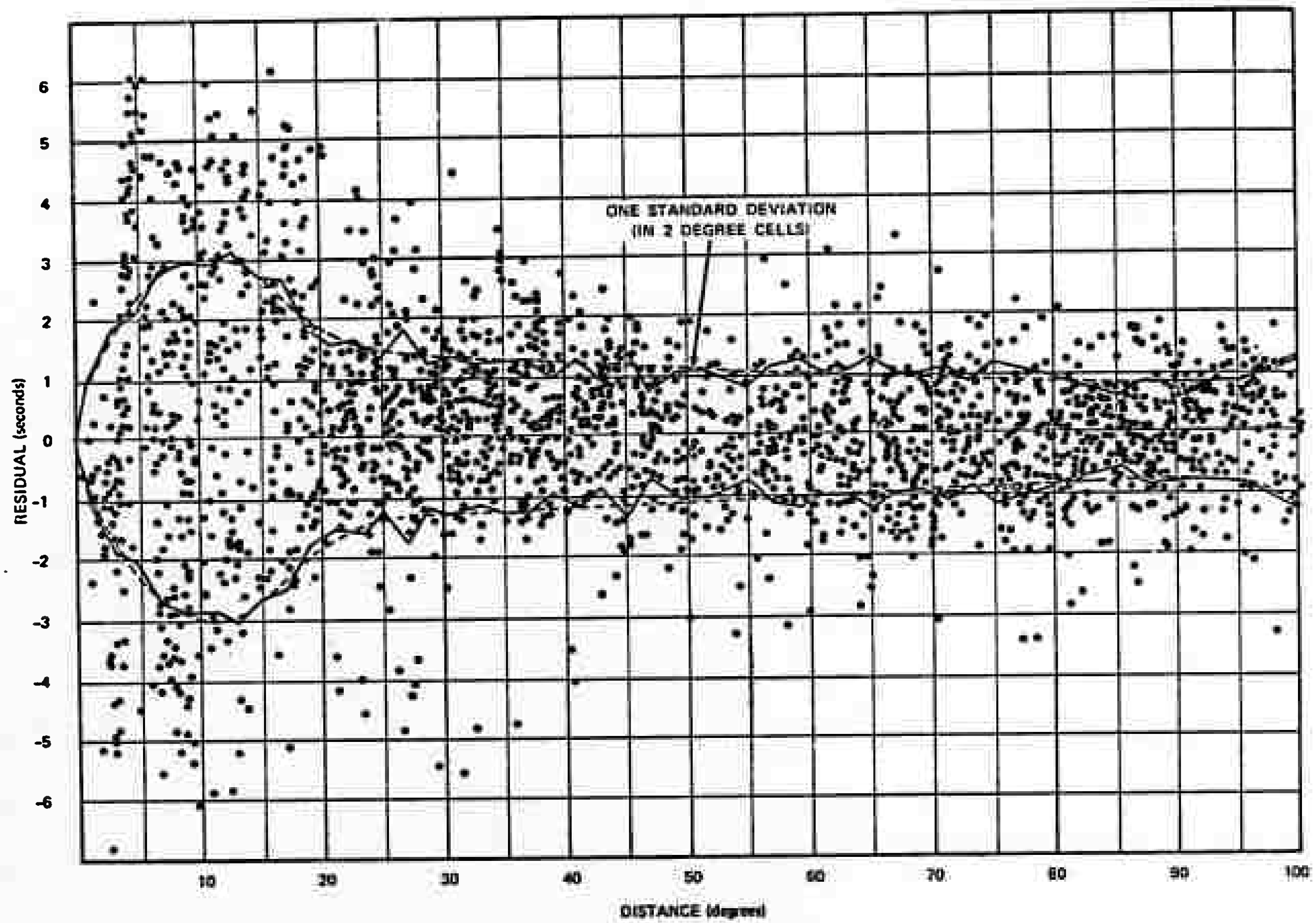
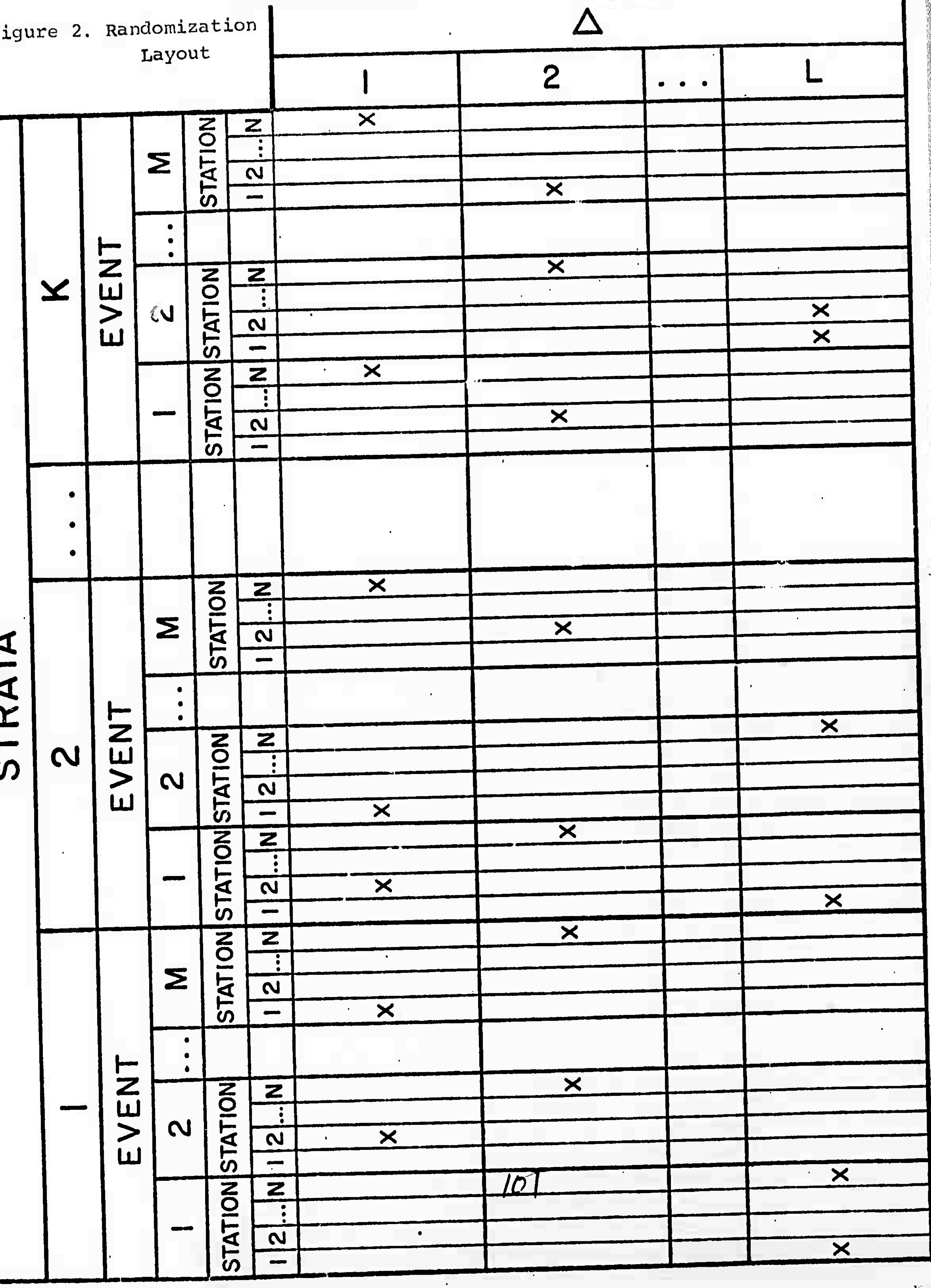


Figure 1. Explosion data residuals

igure 2. Randomization Layout



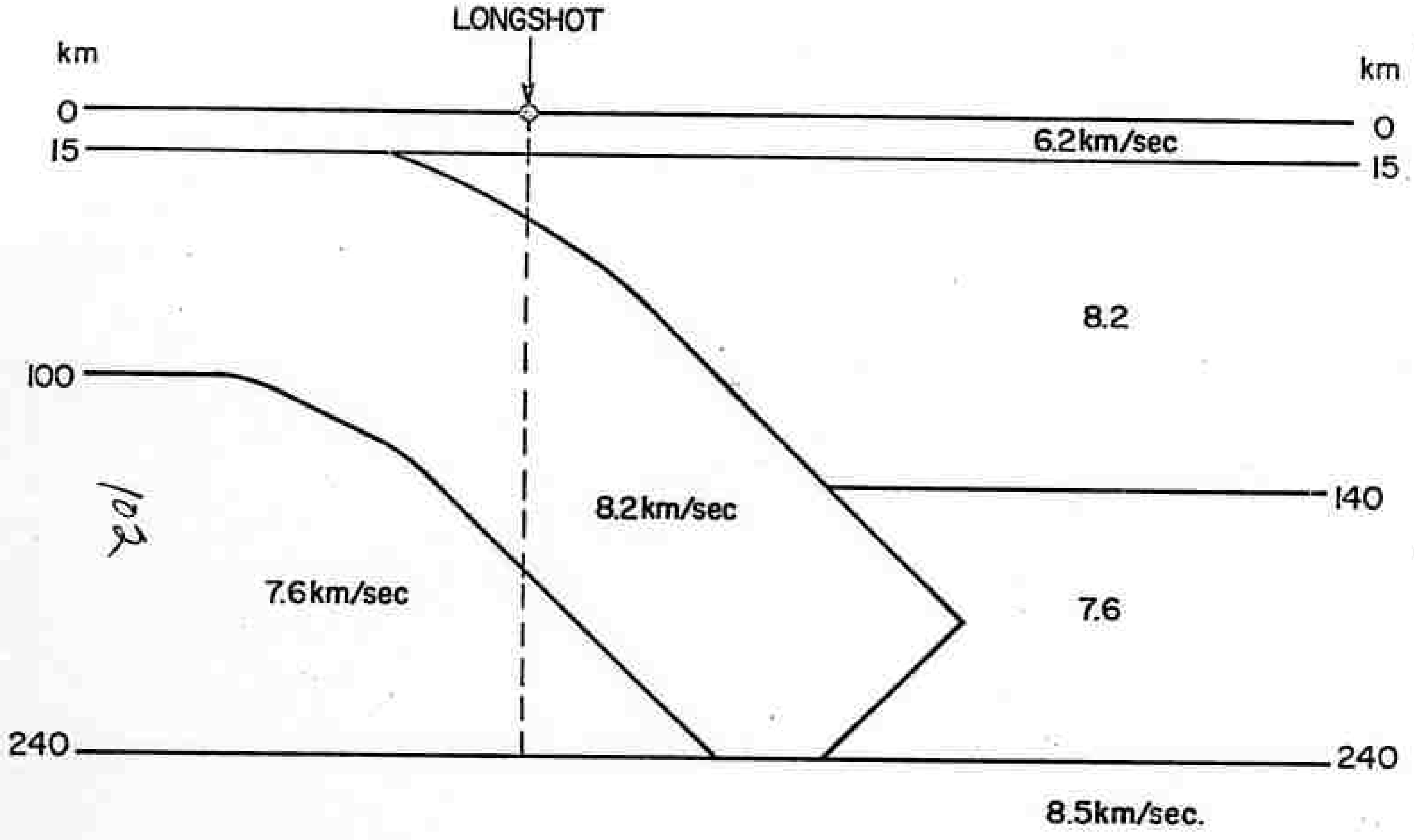
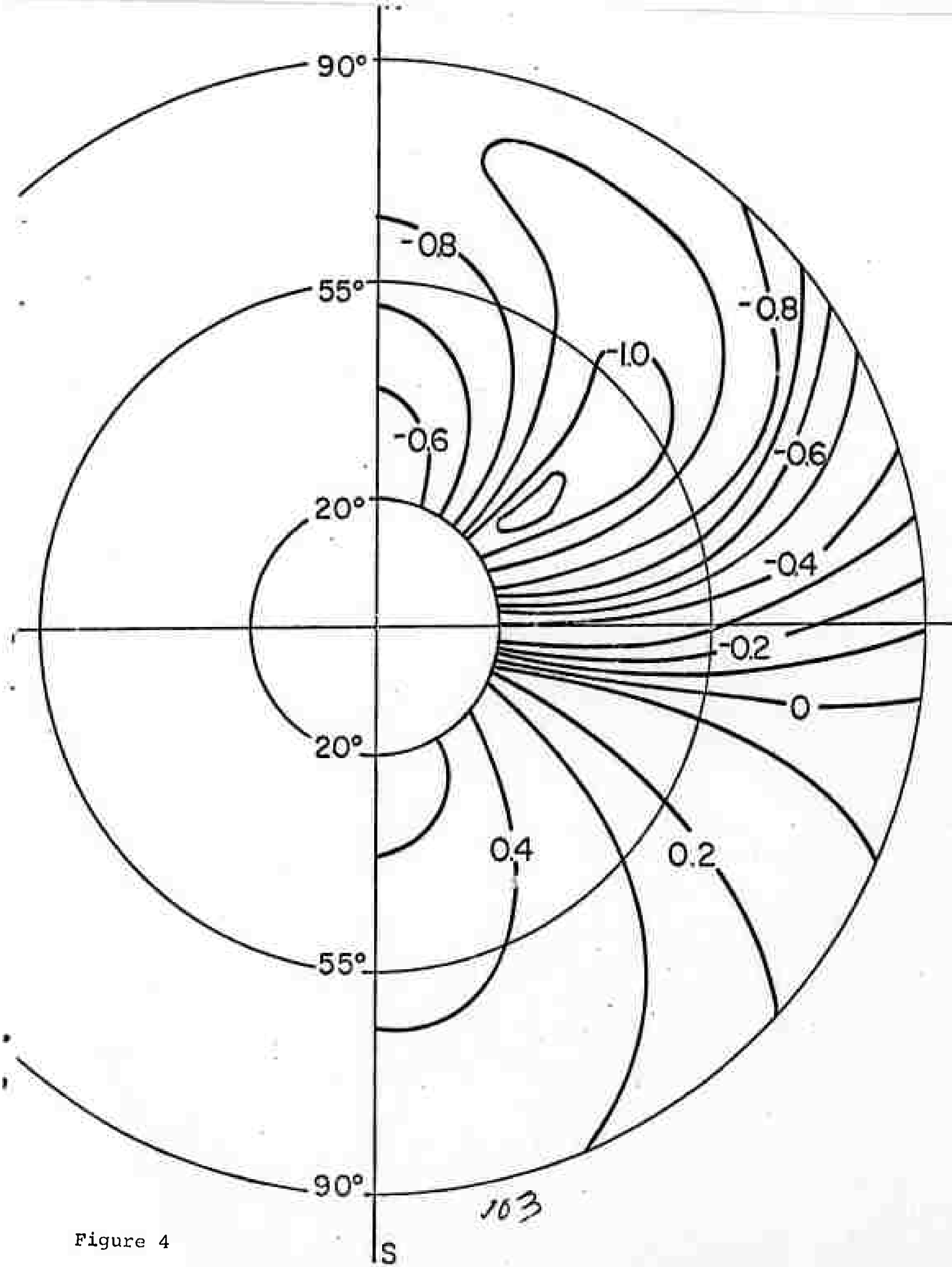


Figure 3 AMCHITKA MODEL



103

Figure 4

COMPUTED TRAVEL TIME ANOMALIES (SEC.)

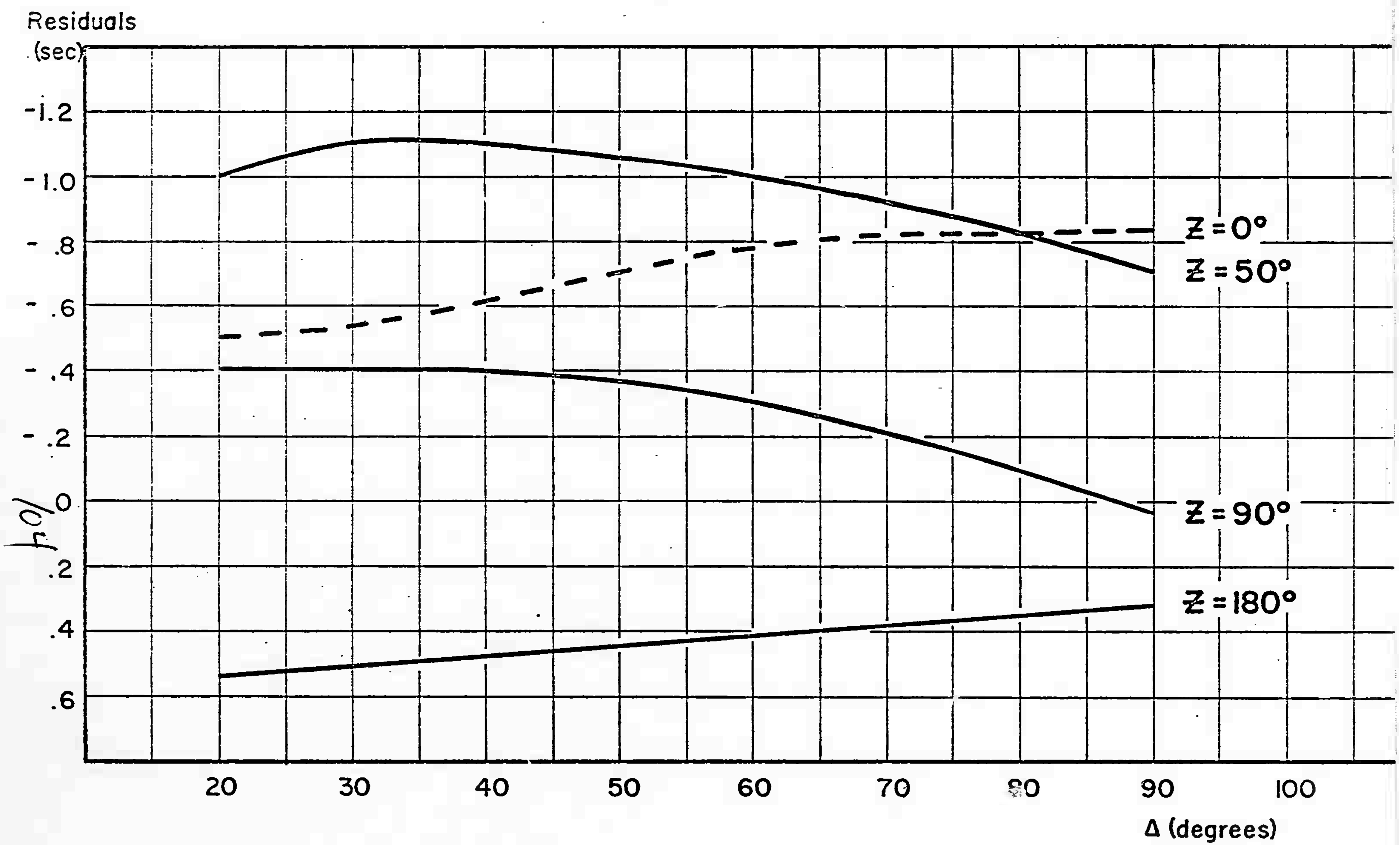
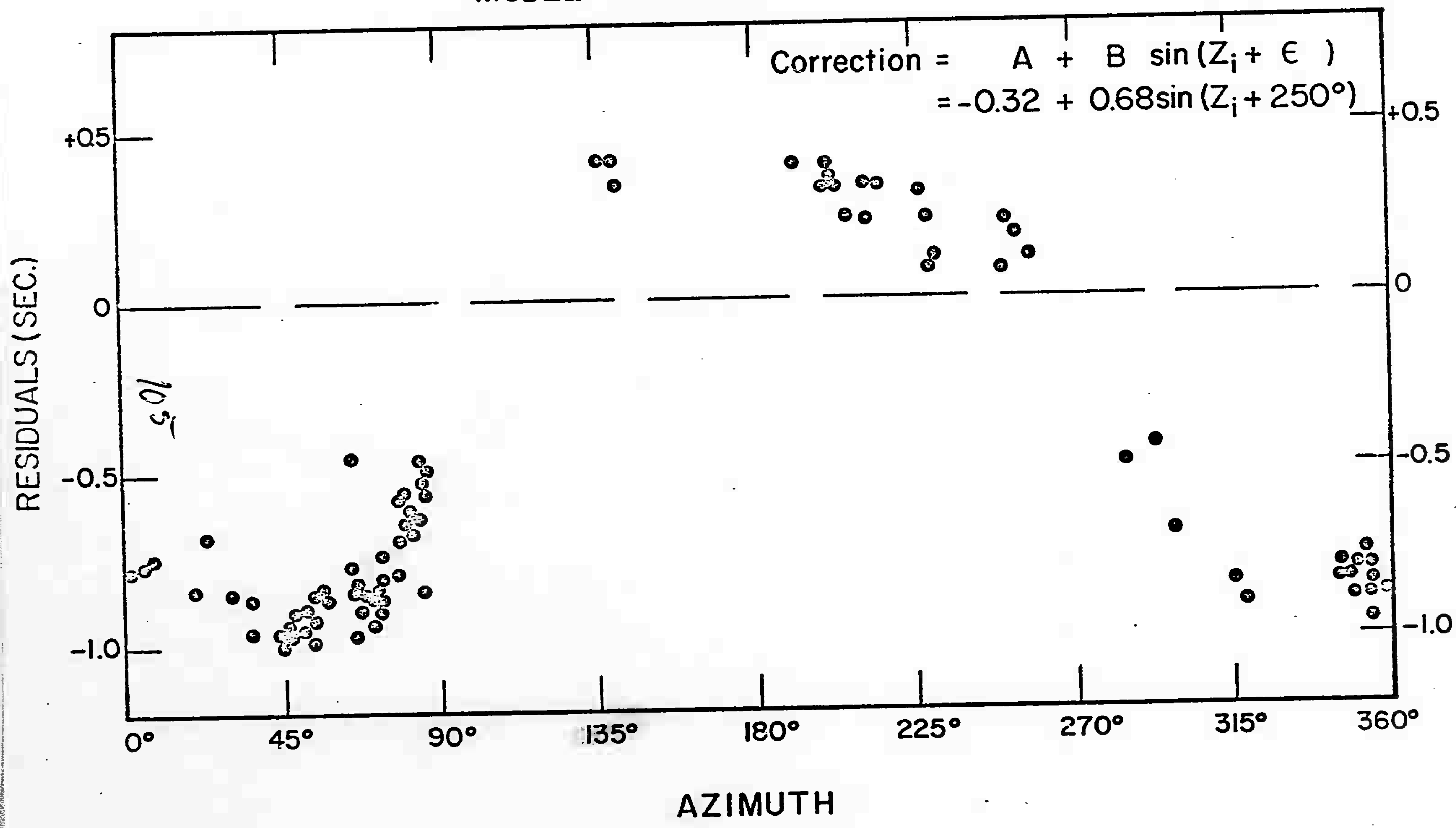


Figure 5 **COMPUTED RESIDUALS**

Figure 6 MODEL RESIDUALS VS. AZIMUTH



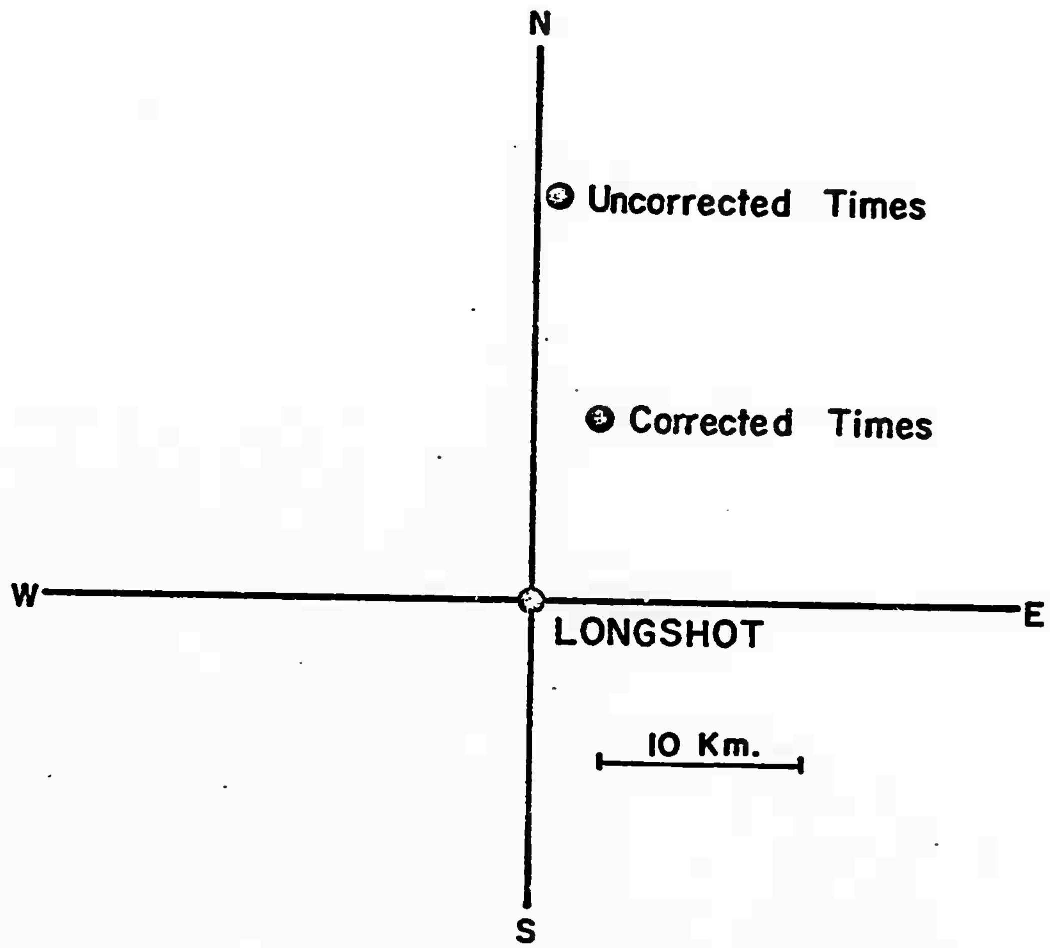


Figure 7 LOCATION ERRORS FOR LONGSHOT

REFERENCES

Bolt, B. and O. Nuttli (1966). P-wave residuals as a function of azimuth, 1, observations, J. Geophys. Res. 71, 130-136.

Cleary, J. and A. Hales (1966a). An analysis of the travel time of P-waves to North American stations, in the distance range 32° to 100°, Bull. Seism. Soc. Am. 56, 467-489.

Cleary, J. and A. Hales (1966b). Azimuthal variation of U.S. station residuals, Nature 210, 619-620.

Cleary J. and K. J. Muirhead (1969). Comparison of the 1968 P tables with times from nuclear explosions (I) Longshot and Greeley, Earth Planet. Sci. Letters 7, 119-124.

Cochran, William G. (1963). Sampling Techniques, Second Edition, John Wiley and Sons, Inc., New York.

Cornfield, J. (1944). On samples from finite populations, J. Amer. Statist. Ass. 39, 236-239.

Dixon, W. J. (1953). Processing data for outliers, Biometrics 9, 74-89.

Douglas, A. (1970). How accurate are the 1968 P tables?, Earth Planet. Sci. Letters 8, 250-252.

Douglas, A. and R. C. Lilwall (1968). Does epicenter bias exist?, Nature 220, 469-270.

Engdahl, E. R., J. N. Taggart, John L. Lobdell, E. P. Arnold and G. E. Clawson (1968). Computational methods, Bull. Seism. Soc. Am. 58, 1339-1344.

Freedman, H. (1966). The 'little variable factor': a statistical discussion of the reading of seismograms, Bull. Seism. Soc. Am. 56, 593-604.

Freedman, H. (1968). Seismological measurements and measurement error, Bull. Seism. Soc. Am. 58, 1261-1271.

Godambe, V. P. (1955). A Unified theory of sampling from finite populations, _J. Royal Stat. Soc., Series B_ 17, 269-278.

Govindarajulu, Z. and R. T. Leslie (1970). Annotated bibliography on robustness studies, _Technical Report No. 7_, Department of Statistics, University of Kentucky.

Graybill, Franklin A. (1961). _An Introduction to Linear Statistical Models, Volume I_, McGraw-Hill Book Company, Inc., New York.

Hanurav, T. V. (1968). Hyper-admissibility and optimum estimators for sampling finite populations, _Ann. Math. Stat._ 39, 621-642.

Herrin E. and G. Sorrells (1969). Travel-time models (Abstract), _Trans. Am. Geophys. Union_ 50, S-54.

Herrin, E. and J. N. Taggart (1968a). Regional variations in P travel times, _Bull. Seism. Soc. Am._ 58, 1325-1337.

Herrin, E. and J. N. Taggart (1968b). Source bias in epicenter determinations, _Bull. Seism. Soc. Am._ 58, 1791-1796.

Herrin, E., W. Tucker, J. N. Taggart, D. W. Gordon and J. L. Lobdell (1968). Estimation of surface-focus P travel times, _Bull. Seism. Soc. Am._ 58, 1273-1291.

Householder, Alston S. (1953). _Principles of Numerical Analysis_, McGraw-Hill Book Company, Inc., New York.

Kempthorne, Oscar (1952). _Design and Analysis of Experiments_, John Wiley and Sons, Inc., New York.

Kempthorne, Oscar (1955). The randomization theory of experimental inference, _J. Amer. Statist. Ass._ 50, 946-969.

Lilwall, R. C. and A. Douglas (1970). Estimation of P-wave travel times using the joint epicentre method, _Geophys., J. R. astr. Soc._ 19, 165-181.

Madow, W. G. (1948). On the limiting distribution of estimates based on samples from finite universes, Ann. Math. Stat. 19, 535-545.

Muirhead, J. J. and J. Cleary (1969). Comparison of the 1968 P tables with times from nuclear explosions (II) the Marshall Islands and Sahara series, Earth Planet. Sci. Letters 7, 132-136.

Nizai, Mansour (1966). Corrections to apparent azimuths and travel-time gradients for a dipping Mohorovici discontinuity, Bull. Seism. Soc. Am. 56, 491-509.

Otsuka, M. (1966). Azimuth and slowness anomalies of seismic waves measured on the central California seismographic array, 1, observations, Bull. Seism. Soc. Am. 56, 223-239.

Ritsema, A. R. (1959). Note on the deviations of P-waves recorded at Djakarta station, Geofisca Pure e, Appl. 43, 159-166.

Tucker, W., E. Herrin and A. W. Freedman (1968). Some statistical aspects of the estimation of seismic travel times, Bull. Seism. Soc. Am. 58, 1243-1260.

Tukey, John W. (1960). A survey of sampling from contaminated distributions, in Contributions to Probability and Statistics, E. Olkin, et al., Stanford University Press, Stanford, 448-485.

Tukey, John W. (1962). The future of data analysis, Ann. Math. Stat. 33, 1-67.

Veith, C. and G. Clawson (1972). Magnitude from short period P-wave data, Bull. Seism. Soc. Am. 62, in press.